

This Page Is Inserted by IFW Operations
and is not a part of the Official Record

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

IMAGES ARE BEST AVAILABLE COPY.

**As rescanning documents *will not* correct images,
please do not report the images to the
Image Problems Mailbox.**

THIS PAGE BLANK (USPTO)

8B05

SEARCH REQUEST FORM

8-70

Requestor's

Name: S. Ugar

Serial

Number: 08/949,904Date: 8/3/98Phone: 305-2181Art Unit: 1642

Search Topic:

Please write a detailed statement of search topic. Describe specifically as possible the subject matter to be searched. Define any terms that may have a special meaning. Give examples or relevant citations, authors, keywords, etc., if known. For sequences, please attach a copy of the sequence. You may include a copy of the broadest and/or most relevant claim(s).

Please Search SEQ ID #1
SEQ ID #2

SEQ ID #3

Reverse Transcribed SEQ ID #2

also Interference Search

Push For Seq ID #1
" " #2
" " #3

*08-848439
08-796153

Reverse Transcribed SEQ ID #2

I need this by COB 8/5

Thanks
Susan

STAFF USE ONLY

Date completed: 8-5-98Searcher: PAE 224291Terminal time: 3Elapsed time: 4:15

CPU time: _____

Total time: _____

Number of Searches: _____

Number of Databases: 11

Search Site

____ STIC

____ CM-1

____ Pre-S

Type of Search

2 N.A. Sequence0 A.A. Sequence

____ Structure

____ Bibliographic

Vendors

✓ IG

____ STN

____ Dialog

____ APS

____ Geninfo

____ SDC

____ DARC/Questel

MPI Other

THIS PAGE BLANK (USPTO)

[M] [P] [E] [H]

(TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

MPsrch_nn n.a. - n.a. database search, using Smith-Waterman algorithm

Run on: Wed Aug 5 04:44:53 1998; MasPar time 2726.89 Seconds
1365.563 Million cell updates/sec

Tabular output not generated.

Title: >US-08-848-439-1
Description: (1-2027) from US08848439.seq
Perfect Score: 2027
N.A. Sequence: 1 GAATTCGGGCTTCATGGCCT.....AAAAAAGAGCGCGCGC 2027
Comp: CTTAAGCCGGAGTACCGGA.....TTTTTTTTTTCGGCGGCG

Scoring table: TABLE default
Gap 6

Nmatch STD : Dbase 0; Query 0

Searched: 532261 seqs, 918536377 bases x 2

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: emb155
1:em_ba 2:em_htg 3:em_hum1 4:em_hum2 5:em_in 6:em_om
7:em_or 8:em_ov 9:em_pat 10:em_pl 11:em_ro
genbank107

Database: 12:gb_ba 13:gb_htg 14:gb_in 15:gb_om 16:gb_ov 17:gb_pat
18:gb_ph 19:gb_pl 20:gb_pr1 21:gb_pr2 22:gb_ro 23:gb_st
24:gb_sts 25:gb_sy 26:gb_un 27:gb_v1

Statistics: Mean 12.042; Variance 6.972; scale 1.727

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES				% Query		Match Length		DB ID		Description		Pred. No.	
Result	No.	Score	Match	Length	DB	ID	Description	Pred. No.					
1	1043	51.5	2028	22	MMU88567		Mus musculus secreted	0.00e+00					
2	993	49.0	2031	22	AF017989		Mus musculus secreted	0.00e+00					
3	959	47.3	1776	22	D50462		Mouse SDF5 mRNA, compl	0.00e+00					
4	765	37.8	882	21	AF017986		Homo sapiens secreted	0.00e+00					
5	110	5.4	1984	21	AF017988		Homo sapiens secreted	3.62e-50					
6	106	5.2	2094	21	AF017987		Homo sapiens secreted	1.75e-47					
7	106	5.2	4240	15	BT085945		Bos taurus frizzled-re	1.75e-47					
8	103	5.1	2075	21	AF001900		Homo sapiens secreted	1.77e-45					
9	103	5.1	4469	21	AF056087		Homo sapiens secreted	1.77e-45					
10	101	5.0	2659	22	MMU88566		Mus musculus secreted	3.80e-44					
11	97	4.8	7218	17	I66494		Sequence 14 from paten	1.72e-41					
12	96	4.7	7218	17	I66494		Sequence 14 from paten	1.78e-41					
13	83	4.1	2075	16	AF006508		Gallus gallus crescent	2.54e-32					
14	60	3.0	2184	21	HSU82169		Human frizzled homolog	7.84e-18					
15	58	2.9	2334	20	HSU43318		Human putative transme	1.28e-16					

16	57	2.8	2328	16	AF031830	Gallus gallus 7-transm	5.11e-16
17	56	2.8	2421	22	MMU43321	Mus musculus putative	2.03e-15
18	51	2.5	1881	22	AF054623	Mus musculus frizzled-	1.84e-12
19	48	2.4	1181	16	AF059570	Xenopus laevis secreta	2.00e-10
20	49	2.4	10772	14	AF012089	Drosophila melanogaste	1.67e-11
21	46	2.3	10772	14	AF012089	Drosophila melanogaste	1.39e-09
22	45	2.2	1912	22	RATFRZH	Rattus norvegicus Bros	5.08e-09
23	42	2.1	1727	16	AF031831	Gallus gallus 7-transm	2.36e-07
24	41	2.0	1923	21	HUMFR12	Human frizzled gene pr	8.32e-07
25	39	1.9	215	17	I28278	Sequence 5 from patent	9.95e-06
26	38	1.9	4540	22	RATFRZH	Rattus norvegicus Bros	3.39e-05
27	31	1.5	215	17	I28278	Sequence 5 from patent	1.18e-01
28	30	1.5	354	14	OFU89259	Oxytricha fallax 57kd	3.53e-01
29	30	1.5	4297	19	ATHSC701	A.thaliana hsc70-1 gen	3.53e-01
30	31	1.5	38262	14	CELK02A2	Caenorhabditis elegans	1.18e-01
31	31	1.5	149435	13	HS111189	Human DNA sequence ***	1.18e-01
32	29	1.4	156	24	DM36D7S	D. melanogaster STS de	1.03e+00
33	29	1.4	307	24	AF010529	Plasmodium falciparum	1.03e+00
34	28	1.4	354	14	OFU89259	Oxytricha fallax 57kd	2.96e+00
35	28	1.4	397	24	GI4542	human STS SHGC-11394	2.96e+00
36	29	1.4	1073	19	ATCOR47	A.thaliana cor47 mRNA	1.03e+00
37	28	1.4	1566	17	E03096	DNA encoding calphobin	2.96e+00
38	29	1.4	2028	21	HSU56813	Human polycystin mRNA	1.03e+00
39	29	1.4	2112	22	MMBK	M. musculus mRNA for EB	1.03e+00
40	29	1.4	2260	22	MMU43320	Mus musculus putative	1.03e+00
41	29	1.4	4798	14	PFSC04076	Plasmodium falciparum	1.03e+00
42	29	1.4	8595	14	DDU14576	Dictyostellium discoide	1.03e+00
43	28	1.4	72698	19	AB007645	Arabidopsis thaliana g	2.96e+00
44	28	1.4	110102	19	ATF28J12	Arabidopsis thaliana D	2.96e+00
45	29	1.4	118710	13	AC004392	*** SEQUENCING IN PROG	1.03e+00

ALIGNMENTS

RESULT 1 MMU88567 2028 bp mRNA ROD 22-APR-1997
LOCUS Mus musculus secreted frizzled related protein sFRP-2 (Sfrp2) mRNA,
DEFINITION complete cds.
ACCESSION U88567
NID 91946342
KEYWORDS house mouse.
SOURCE Mus musculus
ORGANISM Eukaryotae; Mitochondrial eukaryotes; Metazoa; Chordata;
Vertebrata; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae;
Murinae; Mus.
REFERENCE 1 (bases 1 to 2028)
AUTHORS Rattner,A., Hsieh,J.-C., Smallwood,P.M., Gilbert,D.J.,
Copeland,N.G., Jenkins,N.A. and Nathans,J.
TITLE A family of secreted proteins contains homology to the
cysteine-rich ligand-binding domain of frizzled receptors
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 94 (7), 2859-2863 (1997)
MEDLINE 97250455
REFERENCE 2 (bases 1 to 2028)
AUTHORS Rattner,A., Hsieh,J.-C., Smallwood,P.M., Gilbert,D.J.,
Copeland,N.G., Jenkins,N.A. and Nathans,J.
TITLE Direct Submission
JOURNAL Submitted (05-FEB-1997) Molecular Biology and Genetics, HHMI/Johns
Hopkins University School of Medicine, 725 North Wolfe Street, PCTB
805, Baltimore, MD 21205
FEATURES
Location/Qualifiers
1..2028
/organism="Mus musculus"
/db_xref="taxon:10090"
/chromosome="3"
/map="near Fgg"
236..1123
/gene="Sfrp2"
236..1123
/gene="Sfrp2"
/gene="Sfrp2"
/codon_start=1
/product="secreted frizzled related protein sFRP-2"
/db_xref="PID:g1946343"

/translation="MPRGPASILLVLLASHCCLSARGFLFGQDPDSYKRSNCKPIP
ANLQCHIEIQNWRPLNLLHETMKEVLEAGAWIPLYMKOCHPDTRKFLCSFAPV
CLDDDETQPCSHLCVOVKDRCAPVMSAFQPMPLDLECDPNDLQIPLASSDH
LLPATEAPKYCEACKTKNEDNDIMETLCKNDLFAIKVKEIYVNRDRIILETKS
KTIYKLVGSERDLKSLVWLKDSLQCTCEEMNDINAPLYVMGQKSGELVITSKRW
OKGOREFRKRSIRKLQC"
BASE COUNT 469 a 586 c 480 g 492 t 1 others
ORIGIN

Query Match 51.5%; Score 1043; DB 22; Length 2028;
Best Local Similarity 81.8%; Pred. No. 0.00e+00;
Matches 1558; Conservative 0; Mismatches 312; Indels 34; Gaps 24;

Db	127	CCAGAGCCCCACGAGAGAGGAGGAGTCCCGGACAGAGCTCGAGCTCCGGCCGCTCT	186
QY	147	CCCGAGTCCGACCCAGGAGAGAGGCGGCGGAGCAAGCTCGAACTCCCGCGCCTCG	206
Db	187	CGCTTCCCGCCTCGGCTCCCTCGCCCGCGGGTGGCTAGTCCAGAGTCCCGCGGG	246
QY	207	CCCTTCCCGGCTCGGCTCCCTCGCCCGCTCGGGTGGCGGCCAGTGTGCGAGGG	266
Db	247	CCCTGCTCGTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCT	306
QY	267	CCCTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCT	326
Db	307	GCTCTTCTTCTTCCGCGAGCCGACTTCTCTCAAGGCGAGCAACTGCAAGCCCATCCC	366
QY	327	GCTCTTCTTCTTGGCGAGCCGCACTTCTCTCAAGGCGAGCAACTGCAAGCCCATCCC	386
Db	367	CGCAACCTCGAGTGTGCGACGCGATCGAGTACCAAGCAATCGGCTGCCCACTGCT	426
QY	387	GGCAACCTCGAGTGTGCGACGCGATCGAGTACCAAGCAATCGGCTGCCCACTGCT	446
Db	427	GGGCGAGGACCATGAAGAGGTGTGAGAGCGGCGGCGCTGTGATTCGCTGTGCTAT	486
QY	447	GGGCGAGGACCATGAAGAGGTGTGAGAGCGGCGGCGCTGTGATTCGCTGTGCTAT	506
Db	487	GAAGCAGTGCACCCGAGACACCAAGAGTTCTGTGCTGTGCTGTGCTGTGCTGTGCT	546
QY	507	GAAGCAGTGCACCCGAGACACCAAGAGTTCTGTGCTGTGCTGTGCTGTGCTGTGCT	566
Db	547	CGACGACCTAGATGAGACCATCCAGCGGTGTGCTACTGCTGTGCTGTGCTGTGCTGTGCT	606
QY	567	CGATGACCTAGAGAGACCATCCAGCGATGCCACTGCTGTGCTGTGCTGTGCTGTGCTGTGCT	626
Db	607	CTGCGCCCGGCTCATGTCGCGCTTCGCGCTTCGCGCTTCGCGCTTCGCGCTTCGCGCT	666
QY	627	CTGCGCCCGGCTCATGTCGCGCTTCGCGCTTCGCGCTTCGCGCTTCGCGCTTCGCGCT	686
Db	667	TTTCCCGGAGGACAGCTCTGTGATCCGCTTCGCTGTGCTGTGCTGTGCTGTGCTGTGCT	726
QY	687	TTTCCCGGAGGACAGCTCTGTGATCCGCTTCGCTGTGCTGTGCTGTGCTGTGCTGTGCT	746
Db	727	CACAGAGGAAGTCCCAAGGTGTGAGAGCTGTGAGAGCCCAAGATGAGGACGACACCA	786
QY	747	CACGAGGAGTCCCAAGGTGTGAGAGCTGTGAGAGCCCAAGATGAGGACGACACCA	806
Db	787	CATCATGGAACCCCTTTGAAATGACTTCCGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCT	846
QY	807	CATAATGGAACCCCTTTGAAATGACTTTCGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCT	866
Db	847	GTACATCACAGACACCAAGATCATCTGTGAGACAAAGACAGACCATTTTCAAGCT	906
QY	867	GTACATCACAGACACCAAGATCATCTGTGAGACAAAGACAGACCATTTTCAAGCT	926
Db	907	GAACGCGTGTCCGAAAGGAGCTGAAGAAATCCGTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCT	966
QY	927	GAACGCGTGTCCGAAAGGAGCTGAAGAAATCCGTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCT	986
Db	967	GTGACCTGTGAGGAGATGAACGACATCAACGCTCCGCTATCTGTGCTGTGCTGTGCTGTGCTGTGCT	1026
QY	987	GTGACCTGTGAGGAGATGAACGACATCAACGCTCCGCTATCTGTGCTGTGCTGTGCTGTGCTGTGCT	1046

Db	1027	GGCGCGGAGCTGGTGATCACCTCGTGAAACGGTGGCAGAGGCGCCAGAGAGTTCAA	1086
QY	1047	GGGTGGGAGCTGGTGATCACCTCGTGAAACGGTGGCAGAGGCGCCAGAGAGTTCAA	1106
Db	1087	GGCATCTCCCGCAGCATCCGCAAGCTGCAATGCTAGTTCC-CAGTGGGTGGCTCTCT	1145
QY	1107	GGCATCTCCCGCAGCATCCGCAAGCTGCAATGCTAGTTCCCGCGCATCTGATGCTCGGA	1166
Db	1146	CCATCCAGGCGCTGAGCTCTGTAGACCATCTCCGCTCCGGCAGCTCATTTCCGGTTCCTC	1205
QY	1167	CAGGCTGTCTCCAGCAGCAGGCTGACCATTTCTCTCCGGATCTCAGCTCCCTCCCTC	1226
Db	1206	AAGCACAGTCCGGGAAAGCTACAGCCCGAGCTTGGAGCCGCTTCCCTGCTCTCTGATG	1265
QY	1227	AAGCACACTCTAGCT-GCTCCAGCTCTCAGCTGGGAGCTTCCCTGCTCTTTTTCAGC	1285
Db	1266	TGCTATCTCCCTAAGATCTCTGAGTTATAAGGCGCTTAGGAGCTTGGAAACCCATAGCTG	1325
QY	1286	TTTGCATCTCCCGAGCATTTCTGAGTTATAAGGCGCACAGGAG---TGA-----TAGCTG	1336
Db	1326	TTTTCACGGAAGGAAAGCCCATCCAGATCTGTACAAATATTCAAACTATAAATC	1385
QY	1337	TTTTCACCTAAG-GAAAGCCCGACCGAATCTGTAGAAATATTCAAACTATAAATC	1395
Db	1386	ATGACTATTTTATGAGTTTATAGAACAGCTCTGTTTAAAGTTAGTTTGAATAGCTGA	1445
QY	1396	ATGAATATTTTATGAGTTTAAATAGCTCACTTTAAAGCTAGTTTGAATAGCTGA	1455
Db	1446	GTACTTTGACCCGAGGAGCTTCTCTCTTTTGTGCTAGTCTGTTGGCTTATACCGTGCAT	1505
QY	1456	--ACTGTGACTTG-GGTCTGGTTGGTTGT--GTTTGTGTTTGTGATC-AGC-TGATTT	1508
Db	1506	TAGTTTGGCATGTGACGCGAATGTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT	1565
QY	1509	TCACTTCCAC-TGAGGTGTCTAATACATGCAAAATGCTTCAATTTCTCTGTGGCCAA	1567
Db	1566	GCTTGGGCGCCAGACTTAGTTGAGATAAGCTGGCTGTATCTCAAAAGTCTTCTCTCAG	1625
QY	1568	ACTTGGGTGCACAAACCTGTTGAGATAAGCTGGCTGTATCTCAACATCTTCTATCAG	1627
Db	1626	TTCCAGGCTGAGATCGGAGCTTAAGCTTTCAAACATTTCTGTTGCTTTTATGCTCT	1685
QY	1628	CTCCAGAC-TGAGACTCAGTGTCTAAGCTTACAAACATTCATCTTTTATACCTTCAT	1686
Db	1686	CATGAGCTCTGACCATTTGATGCTTCCATCCAGCTACAGAACTTCAAGTTTAAAGCA	1745
QY	1687	GG-GAATTAACCTGTATGATAT-CACATTCAGCTACAAATCTTCCATTTATTAGAA	1744
Db	1746	-CACAGTAACCATTCCTCATTCATGATGCTCCCTCAAATAAAAGTGAATACAGTCTATAA	1804
QY	1745	GCACATTAACCATTTCT-ATAGCATGATTTCTCAAGTAAAGGCAAA-AGA-TATA-AA	1800
Db	1805	ATTGAGGAGTATTTAAGCTTTGTTTAAACATCTTTTAAATTCATTTTAAATCATTTT	1864
QY	1801	TTTTTAATTTGACTTGAG--TACTTTAAGCCTTGTTTTAAAC-CA-TTCTTACTTAACT	1856
Db	1865	TTTTTCAACCTAAATCATTTAGCTTACCTGTATATATACGTAGTTAGTTGACCTGAAAG	1924
QY	1857	TTGCA-AATTAAACCATTTAGCTTACCTGTATATATACGTAGTTAGTTTACCTTAAAG	1915
Db	1925	TGTAAAAATATGCTTTAAGCAGACTGTAAATATTTTCAGATAAACATTTATTTCTTTG	1984
QY	1916	TGTAAAAATATGCTTTAAGCAGACTGTAAATATTTTCAGATAAACATTTATTTCTTTG	1974
Db	1985	TATATAACTTTTACATCTGTTGTGCTATATAAAAAA	2028
QY	1975	TATATAACTTTTACATCTGTTGTGCTATATAAAAAA	2018

RESULT 2

AF017989 2031 bp mRNA ROD 21-SEP-1997

LOCUS Mus musculus secreted apoptosis related protein 1 (Sarpl) mRNA,

com2.let4.cds.

Db 909 GTGGGGTGGCTTCTCTCATCCAGGCGCTGAGCTCTGTAGACCACTTCCGCTCCGCGACC 968
Qy 1151 ATCTGATGGCTCGACAGGCTGCTCCAGACAGCGGTGACCAATTTCTGCTCCGGGATC 1210
Db 969 TCATTTCGGGTTTCCCAAGCACAGTCCGGGAAAGCTACAGCCCGAGCTTGGAGCGGCTTG 1028
Qy 1211 TCAGCTCCGGTTCGCCAAGCACACTCTAGCT-GCTCAGCTCTAGCTTGGCAGCTTCC 1269
Db 1029 CCCTGCCCTCTGCATGTGTATCCCTAACATGCTCTGAGTTATAAGGCCCTTAGGAGGCC 1088
Qy 1270 CCCGCGCTTTTGCAGTTTGCATCCCGAGCATTTCTCAGTTATAAGGCCACAGAG- 1326
Db 1089 TTGAAACCCATAGCTGTTTTCACGGAAGAGGAAAGCCCATCCAGATCTTGTACAATA 1148
Qy 1327 -TGA- - - - -TAGCTGTTTTCAGCTAAG-GAAAGCCACCCGAATCTTGTAGAATA 1378
Db 1149 TTCAACTAATAAATCATGACTATTTTATGAAGTTTATAGACAGCTGTTTAAAGGT 1208
Qy 1379 TTCAACTAATAAATCATGATAATTTTATGAAGTTTAAATAGCTACATTTAAAGCT 1438
Db 1209 AGTTTGAATAGCTGTAGTACTTTGACCCGAGGGGCAATTTCTCTTTTGGTCACTGT 1268
Qy 1439 AGTTTGAATAGGTGCA- -ACTGTGACTTGGCTCTGTTGTTGTTGTTGTT-T-T 1493
Db 1269 TGGCTTATACCGTGACATTAGTTGGCCATGTCAGGCGAATTTGTTCTTTTTTTTTTT 1328
Qy 1494 TGAGTCAG--C-TGATTTTCTATCCAC-TGAGGTTGTCATAACATCAAAATGCTTCA 1549
Db 1329 TTTCCCTCTGTTAGCTTGTGGGTCAGAGCTTAGTTGAGATAAGCTGGCTGTTA 1388
Qy 1550 ATTTTCTGTGGGCCAACTTGTGGGTGCACAAACCTGTTGAGATAAGCTGGCTGTTA 1609
Db 1389 TCTCAAGCTCTCTCAGTCCAGCTCCAGCTCAGAACTGGCATCTAAAGCTTCAAAACATTTTCT 1448
Qy 1610 TCTCAACATCTTTCATCAGCTCCAGCTCAGACTCAGTCTCTAAGCTTACAAACATTCAT 1669
Db 1449 TGCTCGTTTATGCCCTCATGAGCTCTGACCATGTCATGCGTTCCCATCCAGCTACAGA 1508
Qy 1670 CATTTTATACCTTCAATGG-GAACTTAACTGTTACATGTAT-CACATTCAGCTACAA 1727
Db 1509 ACTCAGTTTATAAGCA-CACAGTAACCATTCCTCATGTCATGATGCCCTCAAAATAAAA 1567
Qy 1728 ACTTCCATTTATAGAAGCAGATTAACCATTTCT-ATAGCATGATTTCTCAAGTAAAG 1786
Db 1568 GTGAATACAGTCTATAAATGACAGATATTTAAGCTTTGTTTAAACATCTTTTAATTC 1627
Qy 1787 CAAA-AGA-TATA-AATTTTATATGACTTGAG--TACTTTAGCCTTGTTTAATA-C 1840
Db 1628 AATTTTAAATCATTTTTTTTGCAACTAAATCATATGATGCTTACCTGTATATACGTAG 1687
Qy 1841 A-TTCTTACTTAATTTTGTGA-AATTAACCCATGTTAGCTTACCTGTATATACATAG 1898
Db 1688 TAGTGCCTGGAAGTTGTAANAATTTCTTTAAACGACACTGTAATATTTTCAGAT 1747
Qy 1899 TAGTTACCTTTAAAGTTGTAANAATTTGCTTTAAACCAACACTGTAATATTTTCAGAT 1958
Db 1748 AAACATTAATTTCTTTGATATAAATTT 1776
Qy 1959 AAACATTAATTTCTT-GTATATAAATTT 1986

RESULT 4
LOCUS AF017986 882 bp mRNA PRI 21-SEP-1997
DEFINITION Homo sapiens secreted apoptosis related protein 1 (SARPL) mRNA,
partial cds.
ACCESSION AF017986
NID g2415414
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryotes; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 882)

AUTHORS Melkonyan,H., Chang,W.C., Shapiro,J.P., Mahadevappa,M.,
Fitzpatrick,P.A., Kiefer,M.C., Tomei,D.B. and Umansky,S.R.
TITLE SARPs - a new family of proteins that regulate apoptosis
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 882)
AUTHORS Melkonyan,H., Prochazka,V. and Umansky,S.R.
TITLE Direct Submission
JOURNAL Submitted (11-AUG-1997) LXR Biotechnology, Inc., 1401 Marina Way
South, Richmond, CA 94804, USA
FEATURES
Location/Qualifiers
1..882
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="4"
1..>882
/gene="SARPL"
265..>882
/gene="SARPL"
/note="similar to frizzled-like receptor CRD"
/codon_start=1
/product="secreted apoptosis related protein 1"
/db_xref="PID:g2415415"
/translation="MLQGGPSLLLLFLASHCCCLGSARGLFLFGOPDFSYKRSNCKPIP
ANLQCHGEYONMRLPNLGHETKVELEQACAWIPLVMKQCHPDTKFLCISLFAFV
CLDLDETQPCHSRCVQYKDRCAPYMSAFPMDLECDRFPQDNLDLCITPLASSDHL
PATEAPKVCCEAKNKNDNDNDIMETLCKNDFALKIKRREITINR"
BASE COUNT 167 a 318 c 233 g 163 t 1 others
ORIGIN
Query Match 37.8%; Score 766; DB 21; Length 882;
Best Local Similarity 97.6%; Pred. No. 0.00e+00;
Matches 841; Conservative 0; Mismatches 10; Indels 11; Gaps 4:
Db 29 CTGGCTCATTTCTGCTCCCGGGTCCGAGCCCGCGAGCTGCGCGGGGCTTCAGAGCGC 88
Qy 19 CTAGCTCATTTCTGCTCCCGGGTCCGAGCCCGCGAGCTGCGCGGGGCTTCAGAGCGC 78
Db 89 CTCGCCGCGCTGTCCTCCCGGTGTCCTGCTTCTCCGGCCCCAGCCCGCGGCTGCCAGC 148
Qy 79 CTCGCCGCGCT--CCTCCCGGTGTCCTGCTTCTCCGGCCCCAGCCCGCGGCTGCCAGC 136
Db 149 TTTTCGGGGCCCCAGTTCGACCCAGGAGAGAGCGGGCCCCGGAGCAAGCTCGAATCC 208
Qy 137 TTTTCGGGGCCCCAGTTCGACCCAGGAGAGAGCGGGCCCCGGAGCAAGCTCGAATCC 196
Db 209 GCGCGCTTCGCGCTTAAACAGCTCCG-TCCCTCTACCCCTTAGGGGTGCGCG--CCACGA 265
Qy 197 GCGCGCTTCGCGCTTCCCGGGCTCGCTCCTCTGCCCCCTCGGGGTGCGCGCCACGA 256
Db 266 TGCTGCAGGGCCCTGGGTGCGTGTGCTGCTTCTCCGGCTCGCATGCTGCTGGGCT 325
Qy 257 TGCTGCAGGGCCCTGGGTGCGTGTGCTGCTTCTTCTCCGCTCGCACTGCTGCTGGGCT 316
Db 326 CGGCGCGGGCTTCTTCTTTGGCCAGCCCGAGCTTCTCTACAAGCGGAGCAATTCGA 385
Qy 317 CGGCGCGGGCTTCTTCTTGGCCAGCCCGAGCTTCTCTACAAGCGGAGCAATTCGA 376
Db 386 AGCCCATCCCGGCCAACTGCGAGCTGTGCCAGGCTCGAATACAGAACATCGGGTGC 445
Qy 377 AGCCCATCCCGGCCAACTGCGAGCTGTGCCAGGCTCGAATACAGAACATCGGGTGC 436
Db 446 CCAACCTGCTGGGCCACGAGACCATGAAGAGGTGCTGGAGCAGCGCGGCTTGGATTC 505
Qy 437 CCAACCTGCTGGGCCACGAGACCATGAAGAGGTGCTGGAGCAGCGCGGCTTGGATTC 496
Db 506 CGCTGCTCATGAAGCAGTGCACCCCGGACACCAAGAGTTCTGCTGCTGCTTCGCCC 565
Qy 497 CGCTGCTCATGAAGCAGTGCACCCCGGACACCAAGAGTTCTGCTGCTGCTTCGCCC 556
Db 566 CGCTGCTCATGAGTCTAGACGAGACCATTCAGCCATGCTCGNTGCGTGCAGG 625
Qy 557 CGCTGCTCATGAGTCTAGACGAGACCATTCAGCCATGCTCGCTGCTGCTGCAGG 616

```

Db 626 TGAAGGATCGCTGCGCCCGGTTCATGTCGCCCTTC-----CCTGGCCCGACATGCTTG 679
QY 617 TGAAGGACCGCTGCGCCCGGTTCATGTCGCCCTTCGCCCTTCGCCCTGCGCCGACATGCTTG 676
Db 680 AGTGGACCGCTTTCCCGCAGGACAAAGACCTTTGCGATCCCGCTCGCTAGCAGGACACCC 739
QY 677 AGTGGACCGCTTTCCCGCAGGACAAAGACCTTTGCGATCCCGCTCGCTAGCAGGACACCC 736
Db 740 TCCTGCCAGCAGCAGGAGAGCTCCAAAGTATGTGAAGCTCGCAAAATATAAAATGATG 799
QY 737 TCTGCCAGCAGCAGGAGAGCTCCAAAGTATGTGAAGCTCGCAAAATATAAAATGATG 796
Db 800 ATGACAACGACATAATGGAACCGCTTTGTAATAATGATTTTGCACTGAAATAAAGTGA 859
QY 797 ATGACAACGACATAATGGAACCGCTTTGTAATAATGATTTTGCACTGAAATAAAGTGA 856
Db 860 AGGAGATACTACATCAACCG 881
QY 857 AGGAGATACTACATCAACCG 878

RESULT 5
LOCUS AF017988 1984 bp mRNA PRI 21-SEP-1997
DEFINITION Homo sapiens secreted apoptosis related protein 3 (SARP3) mRNA,
complete cds.
ACCESSION AF017988
NID 92415418
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
REFERENCE
1 (bases 1 to 1984)
AUTHORS Melkonian,H., Chang,W.C., Shapiro,J.P., Mahadevappa,M.,
Fitapatric,P.A., Kiefer,M.C., Tomei,D.L. and Umansky,S.R.
TITLE SARPs - a new family of proteins that regulate apoptosis
JOURNAL Unpublished
REFERENCE
2 (bases 1 to 1984)
AUTHORS Melkonian,H., Prochazka,V. and Umansky,S.R.
TITLE Direct Submission
JOURNAL Submitted (11-AUG-1997) LXR Biotechnology, Inc., 1401 Marina Way
South, Richmond, CA 94804, USA
FEATURES
source
1..1984
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="13"
/tissue_type="pancreas"
1..1984
/gene="SARP3"
216..1169
/gene="SARP3"
/feature="similar to frizzled-like receptor, extracellular
protein"
/codon_start=1
/product="secreted apoptosis related protein 3"
/db_xref="PID:g2415419"
/translation="MRRAAAGVTAALALLGALHWPAPCEYDYVQWAEPLHG
RYSRPPQCLDIPADLPICHTVGYKMRPLPNLLEHSAEYVQKQASWLPILAKRCHS
DTQVFLCSLFAVDLPDPIYPCRSCEAVRACGLMEAYGFPEWEMLHKHFPDND
LCIAVQGLHLPATAPPVTKICACEMHSADGLMOMCSDSFVVKMRIRKTIENGR
KLIGAKKKKLLKPGPLKKRDKTRLVLMKNAGAGCPQLDLSLVMGRKVDGOL
LLMAYVRWDKKNKMKFAVKFMSYPCSLYFFFGAAEPH"
BASE COUNT 353 a 634 c 627 g 370 t
ORIGIN
Query Match 5.4%; Score 110; DB 21; Length 1984;
Best Local Similarity 73.0%; Pred. No. 3.62e-50;
Matches 246; Conservative 0; Mismatches 82; Indels 9; Gaps 3;

Db 380 CATCCCTGCGGACCTGCGGCTGTGCCACAGGTTGGCTACAAAGCGATGCGGCTGCCCAA 439
QY 381 CATCCCGGCCAACCTGCAGCTGTGCCACAGGCTCGAATACCAAGCATGCGGCTGCCCAA 440

```

```

Db 440 CCTGCTGGACGACGAGACCTGCGCGAAGTGAAGCAGCAGCAGCGAGCTGGCTGCCGCT 499
QY 441 CCTGCTGGCCACGAGACCATGAAGAGGTGCTGGAGCAGCGCGCTTGGATCCCGCT 500
Db 500 GCTGGCCAAAGCCTGCGACTCGGATACGAGTCTTCCTGCTGCTGCTCTTTTGGCGCGCT 559
QY 501 GGTCTATGAAGCAGTCCACCGGACACCAAGAAAGTTCCTGCTGCTGCTCTTTCGCCCGCT 560
Db 560 CTGCTCTCGA-----CG-GCCCATCTACCGCTGCGCTGCTGCGAGCGGTGG 610
QY 561 CTGCTCTCGATGACCTAGACGAGACCATCCAGCCATGCCACTCGCTCTGCTGCGAGTGA 620
Db 611 CGCCGCGCTGCGCGCGCTCATGAGGCGCTACGCGCTTCCCTGCGCTGAGATGCTGACTG 670
QY 621 GGACCGCTGCGCGCGCTCATGCTGCGCTTCCGCTTCCCTGCGCGGACATGCTTGAGTG 680
Db 671 CCACAAGTTCCCGCTGGACACGACCTCTGCAATCGCC 707
QY 681 CGACCGTTTCCCGCGGACACGACGCTTTCATCCCG 717

RESULT 6
LOCUS AF017987 2094 bp mRNA PRI 21-SEP-1997
DEFINITION Homo sapiens secreted apoptosis related protein 2 (SARP2) mRNA,
complete cds.
ACCESSION AF017987
NID 92415416
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
REFERENCE
1 (bases 1 to 2094)
AUTHORS Melkonian,H., Prochazka,V., Chang,W.C. and Umansky,S.R.
TITLE SARPs - a new family of proteins that regulate apoptosis
JOURNAL Unpublished
REFERENCE
2 (bases 1 to 2094)
AUTHORS Melkonian,H., Prochazka,V. and Umansky,S.R.
TITLE Direct Submission
JOURNAL Submitted (11-AUG-1997) LXR Biotechnology, Inc., 1401 Marina Way
South, Richmond, CA 94804, USA
FEATURES
source
1..2094
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="8"
/tissue_type="heart"
1..2094
/gene="SARP2"
303..1247
/gene="SARP2"
/feature="similar to frizzled-like receptor, extracellular
protein"
/codon_start=1
/product="secreted apoptosis related protein 2"
/db_xref="PID:g2415417"
/translation="MGIRSGSGRGAALGVLLALGAALLAVGASAEYDVYFOSDIG
PYQSGRYTKPQCVDIPADLRICHNVYKKMVLNLEHETMAEVKQASWVPLLN
KNRAGQVFLCSLFAVCLDRPIYPCRWLCEAVRDSCEPWFQFGYWPMLKCDKF
PEGDVKTAMPFPNTEASKPGTIVCPDNLKSAELIEHLCASEFALRWKIKEVK
EGDKYIVPKKPKLKLGPILKKDLKLVLYLKNAGADPCPHOLDNLSHHFLIMGRKVK
SOYLTAIHKWDKKNKFKFKNFMKMKHNECPTFQSVFK"
BASE COUNT 484 a 598 c 631 g 381 t
ORIGIN
Query Match 5.2%; Score 106; DB 21; Length 2094;
Best Local Similarity 71.8%; Pred. No. 1.75e-47;
Matches 249; Conservative 0; Mismatches 89; Indels 9; Gaps 5;

Db 452 CTCTACACCAAGCAGCCTCAGTGGTGGACATCCCGCGGACCTGCGGCTGTGCCCAA 511
QY 351 CTCTCTACACGCGCAGCATTTGCAAGCCCATCCCGCGCACTGCGAGCTGTGCCCG 410

```

```

Db 512 CGTGGCTACAAGAAGATGTGCTGCCCAACCTGCTGGAGCAGACGACCATGGCGGAGGT 571
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 411 CATGAATACCAAGAACATGGCGCTGCCCAACCTGCTGGCCACGAGACCATGAAGGAGGT 470
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 572 GAACGACGACGACGAGCTGGGTGCCCTGCTCAACAAGAACTGCCACGCCGACCCCA 631
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 471 GCTGGAGCAGCGCGCGCTTGATCCCGCTGGTCAATGAAGCAGTGCACCCGGACACAA 530
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 632 GGTCTCTCTGCTGCTCTTCGCGCGCGCTGCTGCT-G--GACCG-G-C---CCATCTA 682
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 531 GAAGTTCTGTGCTGCTCTTCGCGCGCGCTGCTGCTGATGACCTAGACGAGACCATCCA 590
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 683 CCGCTGCTGCTGCTGCTGCGAGCGCTGCGGACTCTGTCGAGCGGTCATGCAATCTT 742
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 591 GCCATGCCACTGCTGCTGCGTGAAGGACCGTGCGCCCGCTCATGTCGCGCTT 650
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 743 CGGCTTCTACTGCGCGGAGATGCTTAAGTGTGACAAGTTCCCGGAGG 789
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 651 CGGCTTCCCTGGCCGACATGCTTGATGCGACCGTTTCCCGCAGG 697
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 7
LOCUS BTU85945 4240 bp mRNA MAM 21-AUG-1997
DEFINITION Bos taurus frizzled-related protein Frza mRNA, complete cds.
ACCESSION U85945
NID g2337936
KEYWORDS Bos taurus.
SOURCE Bos taurus.
ORGANISM Bos taurus.
REFERENCE 1 (bases 1 to 4240)
AUTHORS Duplaa,C. and D'Amore,P.A.
TITLE Identification and cloning of a novel secreted form of mammalian frizzled: evidence to suggest a role in the control of growth and differentiation
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 4240)
AUTHORS Duplaa,C. and D'Amore,P.A.
TITLE Direct Submission
JOURNAL Submitted (20-JAN-1997) INSERM U441, AV du Haut Leveque, Pessac 33600, France
FEATURES
    source
        1..4240
            /organism="Bos taurus"
            /db_xref="taxon:9913"
            /cell_type="endothelial"
            /tissue_type="heart"
            198..1124
            /note="frizzled related protein"
            /codon_start=1
            /product="Frza"
            /db_xref="PID:g2337937"
            /translation="MGGRWAAAGALIALAAGSASEYDYVSFQSDIGAYQSGR
            FTKPQCVDPADLRCHNVGYKRWLPNLEHETMAEVKQOASWVPLLNKCHIG
            TOVFLSLFAPVCLDRPIYPCRWLCEAVRDSCEPVQFGFYWPEMLKCDKPPGQV
            IATPNATEASPGQTTCVPCDNLKSEAIIEHLCASEFALRMKIKVEKRENGDKK
            IVPKKRKLPLKLPKRLKLLKLVLYLKGADPCDCHLDNLSHHFLIMGRKRVSYLLT
            AIHKWDKKNKEFTFMKMKNHCEPTFSQVFK"
            210..263
            sig_peptide
            misc_feature
            966..968
            /note="encodes N-glycosylation site"
BASE COUNT 1015 a 1073 c 1112 g 1040 t
ORIGIN
Query Match 5.2%; Score 106; DB 15; Length 4240;
Best Local Similarity 71.4%; Pred. No. 1.75e-47;
Matches 252; Conservative 0; Mismatches 92; Indels 9; Gaps 5;

Db 329 CTTCTACCAAGCACCGCAGTCGCTGGACATCCCGGACGACCTGGCGCTGCCACAA 388
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

QY 351 CTTCTCTACAAGCGCAGCAANTTGCAGGCCATCCCGGCCAACCTGCAGCTGTGCCACGG 410
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 389 CGTGGCTACAAGAGGATGCTGCTGCCCAACCTGCTGGAGCAGACGACCATGGCCGAGGT 448
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 411 CATGAATACCAAGAACATGGCGCTGCCCAACCTGCTGGCCACGAGACCATGAAGGAGGT 470
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 449 GAACGACGACGACGAGCTGGGTGCCCTGCTCAACAAGAACTGCCACATCGGCACCCA 508
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 471 GCTGGAGCAGCGCGCGCTTGATCCCGCTGGTCAATGAAGCAGTGCACCCGGACACAA 530
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 509 GGTCTCTCTGCTGCTCTTCGCGCGCGCTGCTGCT-G--GACCG-G-C---CCATCTA 559
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 531 GAAGTTCTGTGCTGCTCTTCGCGCGCGCTGCTGCTGATGACCTAGACGAGACCATCCA 590
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 560 CCGCTGCGCTGCTGCTGCGAGCGCTGCGGACTCTGTCGAGCGGTCATGCAATCTT 619
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 591 GCCATGCCACTGCTGCTGCGTGAAGGACCGTGCGCCCGCTCATGTCGCGCTT 650
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 620 CGGCTTCTACTGCGCGGAGATGCTCAAGTCGACAAAGTTTCCGAGGCGGACG 672
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 651 CGGCTTCCCTGGCCGACATGCTTGATGCGACCGTTTCCCGCAGGACAA 703
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 8
LOCUS AF001900 2075 bp mRNA PRI 25-JUN-1997
DEFINITION Homo sapiens secreted frizzled-related protein mRNA, complete cds.
ACCESSION AF001900
NID g2213818
KEYWORDS human.
SOURCE Homo sapiens.
ORGANISM Homo sapiens.
REFERENCE 1 (bases 1 to 2075)
AUTHORS Finch,P.W., He,X., Kelley,M.J., Uren,A., Schaudties,R.P., Popescu,N.C., Rudikoff,S., Aaronson,S.A., Varmus,H.E. and Rubin,J.S.
TITLE Purification and Molecular Cloning of a Secreted, Frizzled-Related Antagonist of Wnt Action
JOURNAL Proc. Natl. Acad. Sci. U.S.A. (1997) In press
REFERENCE 2 (bases 1 to 2075)
AUTHORS Finch,P.W., Aaronson,S.A. and Rubin,J.S.
TITLE Direct Submission
JOURNAL Submitted (30-APR-1997) LCMB, DBS/NCI, 9000 Rockville Pike, Bethesda, MD 20892, USA
REFERENCE 3 (bases 1 to 2075)
AUTHORS Finch,P.W., Aaronson,S.A. and Rubin,J.S.
TITLE Direct Submission
JOURNAL Submitted (09-JUN-1997) LCMB, DBS/NCI, 9000 Rockville Pike, Bethesda, MD 20892, USA
REMARK Sequence update by submitter
FEATURES
    Location/Qualifiers
        1..2075
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            /chromosome="8"
            /map="8p11.1-p12"
            /clone="HSL"
            /dev_stage="embryonic"
            /tissue_type="lung"
            /cell_line="M426"
            /cell_type="fibroblast"
            303..1244
            /note="FRP"
            /codon_start=1
            /product="secreted frizzled-related protein"
            /db_xref="PID:g2213819"
            /translation="MGIGRSEGGRRGALGVLLAALGASASEYDYVSFQSDIGP
            YQSGREYTKPQCVDPADLRCHNVGYKRWLPNLEHETMAEVKQOASWVPLLNK
            NCHAGTQVFLCSLFAVCLDRPIYPCRWLCEAVRDSCEPVQFGFYWPEMLKCDKFP
            EDGVCIAMTPPNATEASKPGQTTCVPCDNLKSEAIIEHLCASEFALRMKIKVEKKE
            NGDKKIVPKKPKLPKLPKRLKLLKLVLYLKGADPCDCHLDNLSHHFLIMGRKRVKS

```


BASE COUNT	473 a	596 c	626 g	380 t	
ORIGIN					QVLLTAIHKWDKKNEEFNFKMKNKHCPPTQSFVK"
Query Match	5.1%	Score 103;	DB 21;	Length 2075;	
Best Local Similarity	73.5%;	Pred. No. 1.77e-45;			
Matches	151;	Conservative	0;	Mismatches 58;	Indels 0; Gaps 0;
Db	449	CTTCTACACCAACCACTTCAGTTCGTGGACATCCCGCGGACCTCGCGCTGTGCCAAC	508		
Qy	351	CTTCTCTCTACAAGCGCAGCAATTGCAAGCCCATCCCGCCACCTGCGACTGTGCCAGG	410		
Db	509	CGTGGGCTACAGAAGATGCTGCTGCCCAACCTGTGGAGCAGACAGACCATGGCGGAGGT	568		
Qy	411	CATCGAATACAGAATCGGCTGCCCAACCTGTGGCCACGAGACCATGAAGGAGGT	470		
Db	569	GAAGCAGCAGCGCAGCAGCTGGTGCGCCCTGCTCAACAGAACTGCCACGCGGGAGCCCA	628		
Qy	471	GCTGGACGAGCCGGCGCTTGGATCCCGCTGGTCATGAAGCAGTGCCACCGGACACCA	530		
Db	629	GGTCTTCTCTGCTGCTCTTCGCGCGCCGTCTGCTCGTGA	667		
Qy	531	GAAGTCTCTGCTGCTCTCTGCGCCCGCTGCTGCTCGA	569		
RESULT	9				
LOCUS	AF056087	4469 bp	mRNA	PRI	08-APR-1998
DEFINITION	Homo sapiens secreted frizzled related protein mRNA, complete cds.				
ACCESSION	AF056087				
NID	9303350				
KEYWORDS					
SOURCE	human.				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;				
AUTHORS	1 (bases 1 to 4469)				
TITLE	Zhou, Z. and Wang, J.				
JOURNAL	Upregulation of human secreted Frizzled homologue in apoptosis and its down regulation in breast tumors				
REFERENCE	2 (bases 1 to 4469)				
AUTHORS	Zhou, Z. and Wang, J.				
TITLE	Direct Submission				
JOURNAL	Submitted (29-MAR-1998) Radiohemmet Research Laboratory, CCK, Karolinska Institute, Solna, Stockholm S-171 76, Sweden				
FEATURES	Location/Qualifiers				
source	1..4469				
	/organism="Homo sapiens"				
	/db_xref="taxon:9606"				
CDS	303..1244				
	/function="putative Wnt antagonist"				
	/note="FRP; frizzled homolog; cell cycle and apoptosis related; down regulated in breast tumorigenesis"				
	/codon_start=1				
	/product="secreted frizzled related protein"				
	/db_xref="PID:g303351"				
	/translation="MGICRSEGGREGALGVLLALGAALLAVGSAEYDVVFSQDSIGYQSGRFTKPPQCVDDIPADLRKCHNRYKMWLPNLEHETMAEVKQQAQSVPLLN				
	NCHAGTVQVLCSLFAPVCLDRPIYCPWLCEAVRDSCEPVMQFFGYWPEMLKCDKFT				
	EGDYCIAMTPNATEASKPGDTGTPPCDNEALRSEAIIEHLCASEFALRMKIKVKKK				
	NGDKIVPKKKPLKPLKPLKLVLYLKGADPCPCQDLDNLSHFLIMGRKVK				
	QVLLTAIHKWDKKNEEFNFKMKNKHCPPTQSFVK"				
BASE COUNT	1121 a	1097 c	1120 g	1131 t	
ORIGIN					QVLLTAIHKWDKKNEEFNFKMKNKHCPPTQSFVK"
Query Match	5.1%	Score 103;	DB 21;	Length 4469;	
Best Local Similarity	73.5%;	Pred. No. 1.77e-45;			
Matches	161;	Conservative	0;	Mismatches 58;	Indels 0; Gaps 0;
Db	449	CTTCTACACCAACCACTTCAGTTCGTGGACATCCCGCGGACCTCGCGCTGTGCCAAC	508		
Qy	351	CTTCTCTCTACAAGCGCAGCAATTGCAAGCCCATCCCGCCACCTGCGACTGTGCCAGG	410		

BASE COUNT

ORIGIN

Query Match 4.1%; Score 83; DB 16; Length 2075;
Best Local Similarity 71.1%; Pred. No. 2.54e-32;
Matches 140; Conservative 0; Mismatches 57; Indels

Db	152	TGACACGCCCATCCGGCGCAGCATGGCCCTGTGCTATGACATCGGTACTCGGAGATGAGG	211
Qy	373	TGCAAGCCCATCCCGGCCAACCTGCACTGTGCGCAGGCATCGAATACCAAGACATGCGG	432
Db	212	ATCCCAACCTGCTGGAGCACGAGACCATGCCGGAGGTGATCCAGCAGTCTCCACGCTGG	271
Qy	433	CTGCCCAACCTGCTGGGCCACGAGACCATGAGAGAGTGCTGGAGCAGCGCCGCGCTTGG	492
Db	272	CTGCCCTTGTGTGCCCAGGAGTGCCCATCCGAGCGCCAGGATTTCCTCTGCTCCCTCTTC	331
Qy	493	ATCCCGCTGTGTATGAAGCAGTGCACCCGGACACCAAGAAGTTCTCTGTGCTGCTCTTC	552
Db	332	GCGCCCATCTGCCTGGA	348
Qy	553	GCCCGCTGTGCCTGA	569

RESULT	14	
LOCUS	HSU82169	2184 bp mRNA
DEFINITION	Human frizzled homolog (FZD3) mRNA, complete cds.	PR
ACCESSION	U82169	
NID	g1906597	
KEYWORDS		
SOURCE	human.	
ORGANISM	Homo sapiens	
	Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata; vertebrata; Eutheria; Primates; Catarrhini; Homnidae; Homo.	
REFERENCE	1 (bases 1 to 2184)	
AUTHORS	Wang, Y.-K., Samos, C.H., Peoples, R., Perez-Jurado, L.A., Nusse, R. and Francke, U.	
TITLE	A novel human homologue of the Drosophila frizzled wnt receptor gene binds wingless protein and is in the Williams syndrome deletion at 7q11.23	
JOURNAL	Hum. Mol. Genet. 6 (3), 455-472 (1997)	
MEDLINE	97227293	
REFERENCE	2 (bases 1 to 2184)	
AUTHORS	Wang, Y.-K., Peoples, R., Perez-Jurado, L.A. and Francke, U.	
TITLE	Direct Submission	
JOURNAL	Submitted (12-DEC-1996) Howard Hughes Medical Institute, Stanford Medical Center, Beckman Center B201, Stanford, CA 94305, USA	
FEATURES	Location/Qualifiers	
	1..2184	

BASE COUNT

RPPPAKPTLPFPAPASGECGAGGPFVCKREPFVPIILKESHPLYNKVRTQVPNC
AVCYQPSFADERTFATWIGLWSVLCFISTITVATFLIDMDTFRYPRIIFLSA
CYLCVSLGFLVRLVGHASVACSRHNHIEYETGPALCTIVFLLVFFGMASSIMWV
ILSLTWFLAAMKWEAGYGOYFHLAAMLIPSVKSLTALALSSVDGDPVAGICVY
GNQNLNLRBFVLCPLVLYLLVGLFLLAGFVSLFRIRSVIKOGGKTDXLEKLMIRI
GIFLLTVFASIVVACLYEQHYRESWEALACACPGHDTGQPRAKPEYWLMLRYF
MCLVVGTSGVWINSKRTVESRRFTSRCCCRPRRHKGSGMAAGDYPEASAAALTGR
TGPPGPAATYHKQVSLSHV"

BASE COUNT 356 a 803 c 736 g 439 t
ORIGIN

Query Match 2.9%; Score 58; DB 20; Length 2334;
Best Local Similarity 59.9%; Pred. No. 1.28e-16;
Matches 175; Conservative 0; Mismatches 117; Indels 0; Gaps 0;

Db	439	TGTGCGCGGCGATCGGCTACAGCTGACGCACATGCCCAACAGTTCAACACGACGACGC	498
Qy	401		460
Db	499	AGGACGAGCGGGCTGGAGGTGCACCAAGTTCTGGCCGCTGGTGAGATCCAATGCTCGC	558
Qy	461		520
Db	559	CGGACCTCGGCTTCTTCTATGACTATGTACAGCCCATCTGTGCCCCGACTACACACA	618
Qy	521		580
Db	619	AGCGCTGCCGCTGCGGCTGCGGTGCGAGCGCGCCAAAGCCGCGTCTCGCGCTGA	678
Qy	581		640
Db	679	TGCGCCAGTACGGCTTCGCTTGGCCCGAGCGCATGAGCTGGAGCGCGCTCC	730
Qy	641		692

Search completed: Wed Aug 5 06:19:43 1998
Job time : 5690 secs.

THIS PAGE BLANK (10/27)

W P S R E H

(TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

MPsrch_nn n.a. - n.a. database search, using Smith-Waterman algorithm
Run on: Wed Aug 5 07:34:48 1998; MasPar time 322.83 Seconds
854.249 Million cell updates/sec
Tabular output not generated.

Title: >US-08-848-439-1
Description: (1-2027) from US08848439.seq
Perfect Score: 2027
N.A. Sequence: 1 GAATTGGCGCTTCATGGCT.....AAAAAAGGCGCGC 2027
Comp: CTTAAGCGGGAAGTACCGGA.....TTTTTTTTTCCCGCGC

Scoring table: TABLE default
Gap 6

Nmatch STD : Dbase 0; Query 0
Searched: 188442 seqs, 58026449 bases x 2

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: n-geneseq32
1:part1 2:part2 3:part3 4:part4 5:part5 6:part6 7:part7
8:part8 9:part9 10:part10 11:part11 12:part12 13:part13
14:part14 15:part15 16:part16 17:part17 18:part18
19:part19 20:part20 21:part21 22:part22 23:part23
24:part24 25:part25 26:part26 27:part27 28:part28
29:part29 30:part30 31:part31 32:part32 33:part33
34:part34 35:part35 36:part36 37:part37 38:part38
39:part39 40:part40

Statistics: Mean 9.846; Variance 6.883; scale 1.431
Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

Result		Query		Match		Length		DB		ID		Description		Pred. No.	
No.	Score	Match	Length	DB	ID	Description	Pred. No.	No.	Score	Match	Length	DB	ID	Description	Pred. No.
1	201	9.9	204	20	T23535	Human gene signature	1.01e-97	1	201	9.9	204	20	T23535	Human gene signature	1.01e-97
2	58	2.9	2334	37	T89889	Human frizzled gene 5	6.84e-16	2	58	2.9	2334	37	T89889	Human frizzled gene 5	6.84e-16
3	56	2.8	2421	37	T89892	Mouse frizzled gene 8	7.60e-15	3	56	2.8	2421	37	T89892	Mouse frizzled gene 8	7.60e-15
4	44	2.2	91	9	Q51746	Oligonucleotide probe	9.73e-09	4	44	2.2	91	9	Q51746	Oligonucleotide probe	9.73e-09
5	45	2.2	172	32	T76363	Human interleukin 8 a	3.11e-09	5	45	2.2	172	32	T76363	Human interleukin 8 a	3.11e-09
6	43	2.1	204	1	N81164	Base substituted E.co	3.03e-08	6	43	2.1	204	1	N81164	Base substituted E.co	3.03e-08
7	41	2.0	91	9	Q51746	Oligonucleotide probe	2.88e-07	7	41	2.0	91	9	Q51746	Oligonucleotide probe	2.88e-07
8	40	2.0	100	32	T76186	Human IL4 receptor an	8.76e-07	8	40	2.0	100	32	T76186	Human IL4 receptor an	8.76e-07
9	40	2.0	114	12	Q70466	Generic DNA sequence	8.76e-07	9	40	2.0	114	12	Q70466	Generic DNA sequence	8.76e-07
10	40	2.0	114	12	Q70468	Generic DNA sequence	8.76e-07	10	40	2.0	114	12	Q70468	Generic DNA sequence	8.76e-07
11	40	2.0	204	1	N81164	Base substituted E.co	8.76e-07	11	40	2.0	204	1	N81164	Base substituted E.co	8.76e-07
12	39	1.9	114	12	Q70469	Generic DNA sequence	2.65e-06	12	39	1.9	114	12	Q70469	Generic DNA sequence	2.65e-06
13	38	1.9	114	12	Q70467	Generic DNA sequence	7.95e-06	13	38	1.9	114	12	Q70467	Generic DNA sequence	7.95e-06

14	38	1.9	114	12	Q70465	Generic DNA sequence	7.95e-06
15	39	1.9	114	12	Q70469	Generic DNA sequence	2.65e-06
16	38	1.9	114	12	Q70468	Generic DNA sequence	7.95e-06
17	38	1.9	114	12	Q70466	Generic DNA sequence	7.95e-06
18	39	1.9	178	32	T76405	Human endothelin-1 an	2.65e-06
19	39	1.9	190	32	T76452	Chymase antisense oli	2.65e-06
20	36	1.8	114	12	Q70470	Generic DNA sequence	6.96e-05
21	36	1.8	114	12	Q70472	Generic DNA sequence	6.96e-05
22	36	1.8	114	12	Q70473	Generic DNA sequence	6.96e-05
23	34	1.7	114	12	Q70471	Generic DNA sequence	5.87e-04
24	34	1.7	114	12	Q70465	Generic DNA sequence	5.87e-04
25	34	1.7	114	12	Q70471	Generic DNA sequence	5.87e-04
26	34	1.7	114	12	Q70470	Generic DNA sequence	5.87e-04
27	34	1.7	114	12	Q70470	Generic DNA sequence	5.87e-04
28	34	1.7	250	32	T76438	Substance P antisense	5.87e-04
29	35	1.7	317	32	T76274	Human neutrophil elas	2.03e-04
30	32	1.6	114	12	Q70472	Generic DNA sequence	4.71e-03
31	32	1.6	178	32	T76405	Human endothelin-1 an	4.71e-03
32	32	1.6	264	32	T76445	Substance P receptor	4.71e-03
33	30	1.5	88	32	T76170	Human IL3 receptor an	3.65e-02
34	31	1.5	168	32	T76270	Human MDNCF antisense	1.32e-02
35	31	1.5	190	32	T76452	Chymase antisense oli	1.32e-02
36	30	1.5	317	32	T76274	Human neutrophil elas	3.65e-02
37	30	1.5	565	6	Q35072	RCV envelope region n	3.65e-02
38	30	1.5	3871	2	N71302	HSV-1 gB and surround	3.65e-02
39	29	1.4	70	32	T76184	Human IL4 antisense o	9.94e-02
40	29	1.4	128	32	T76233	Human IL6 antisense o	9.94e-02
41	28	1.4	162	32	T76307	Human RANTES antisens	2.67e-01
42	29	1.4	172	32	T76363	Human interleukin 8 a	9.94e-02
43	28	1.4	1466	2	N91353	Vascular anti-coagula	2.67e-01
44	29	1.4	1840	11	Q56611	Feline zona pellucida	9.94e-02
45	29	1.4	2259	37	T89891	Mouse frizzled gene 7	9.94e-02

ALIGNMENTS

RESULT ...1
ID T23535 standard; cDNA to mRNA; 204 BP.
AC T23535
AT 02-SEP-1996 (first entry)
DE Human gene signature HUMGS05382.
KW Gene signature; messenger RNA; mRNA; relative abundance; frequency;
KW human; cloning; mapping; non-biased library; diagnosis; detection;
KW cell typing; abnormal cell function; ss.
OS Homo sapiens.
PN W09514772-A1.
PD 01-JUN-1995.
PF 11-NOV-1994; J01916.
PR 12-NOV-1993; JP-355504.
PA (MATS/) MATSUBARA K.
PA (OKUB/) OKUBO K.
PI Matsubara K. Okubo K;
DR WPI: 95-206931/27.
PT Identifying gene signatures in 3'-directed human cDNA library - e.g.
PT for diagnosis of abnormal cell function, by preparing cDNA that
PT reflects relative abundance of corresp. mRNA in specific human
PT tissues
PS Claim 1; Page 1393; 2245pp; Japanese.
CC A single-stranded DNA (or its complementary strand or the corresp.
CC double-stranded DNA) which comprises one of the 7837 "GS" sequences
CC given in T19001-T26837 and which is able to hybridise to part of
CC human genomic DNA, cDNA or mRNA is claimed. The GS (Gene Signature)
CC sequences were obtained from 3'-directed cDNA libraries prepared
CC from various human tissues; synthesis of cDNA was initiated from the
CC 3'-end of mRNA by using poly(T) as the sole primer. Since the 3'-
CC untranslated sequence is unique to a particular mRNA species, almost
CC all the 3'-oriented cDNAs hybridise with specific mRNAs. Each library
CC is constructed so as to reflect accurately the relative abundance of
CC different mRNAs in the particular tissue from which it was derived.
CC The appearance frequency of a given GS in a cDNA library can be
CC determined (esp. using primers and probes derived from the GS
CC sequences) as a means of diagnosing abnormal cell function or for
CC recognising different cell types.

SQ	Sequence	204 BP;	54 A;	59 C;	35 G;	55 T;
	Query Match	9.9%;	Score 201;	DB 20;	Length 204;	
	Best Local Similarity	99.5%;	Pred. No. 1.01e-97;			
	Matches	201;	Conservative	0;	Mismatches	1; Indels 0; Gaps 0;
Df	1	gattcagctcccggttcccccaagcacactcttagctgcttcagcttcagctgggcagct	60			
Qy	1207	GATCTAGTCCCGTCGCCAAGCACACTCCTAGCTGCTCCAGTCTCAGGTGGGCAGCT	1266			
Df	61	tcccctgccttttgacgctttgatccaccagcttccctgagttataaaggccacagag	120			
Qy	1267	TCCCCCTGCCTTTTGACGTTTGATCCCAGCATTTCCCTGAGTTATAAGCCACAGAG	1326			
Df	121	tggatagctgttttacctaaagaaaagcccccgaattcttgtagaatatattcaaac	180			
Qy	1327	TGGATAGTGTTTTACGTAAGAAGAAAGCCACCAGCAATCTTTAGAAAATTATCAACT	1386			
Df	181	antaaaaatcatgaatatittta	202			
Qy	1387	AATAAATCATGAATATTTTA	1408			
RESULT	2					
ID	T89889	standard; DNA; 2334 BP.				
AC	T89889;					
DE	27-APR-1998	(first entry)				
DE	Human frizzled gene 5 (Hfz5) encoding a Wnt receptor.					
KW	Wnt receptor; human frizzled gene 5; Hfz5 gene;					
KW	signal transduction; cancer; cell growth; cell proliferation;					
KW	mammary tumour; oncogene; therapy; ds.					
OS	Homo sapiens.					
FH	Key	Location/Qualifiers				
FT	CDS	321..2078				
FT		/*tag= a				
PN	WO9739357-A1.					
PD	23-OCT-1997.					
PF	11-APR-1997; U06049.					
PR	12-APR-1996; US-015307.					
PA	(UJO) UNIV JOHNS HOPKINS.					
PA	(STRD) UNIV LELAND STANFORD JUNIOR.					
PI	Andrew D, Bhanot P, Brink M, Hsieh J, Nathans J,					
PI	Nusse R, Samos CH, Wangy;					
DR	WPI; 97-526631/48.					
DR	P-PSDB; W31271.					
PT	Identification of Wnt receptor binding modulators - useful for					
PT	treatment of cancer and growth, development or proliferation related					
PT	disorders					
PS	Disclosure; Page 36-37; 6lpp; English.					
CC	This sequence comprises novel human frizzled gene 5 (Hfz5)					
CC	that encodes a transmembrane receptor, frizzled-5 (see W31771), a					
CC	Wnt receptor (Wntr). Novel frizzled family members have been					
CC	identified in Drosophila, mouse, human and Caenorhabditis (see					
CC	T89885-92) and are considered to be Wnt receptors. Wnt receptors					
CC	can be used in a novel, claimed method of screening for compounds					
CC	which modulate the binding of a Wnt polypeptide (secreted proteins					
CC	involved in cell-to-cell signalling) to a Wnt receptor. Wnt is					
CC	involved in (mammary) cancer and other processes involving growth,					
CC	development and proliferation (both normal and abnormal).					
CC	Modulators identified by the claimed method are useful for					
CC	treatment of diseases related to these conditions.					
SQ	Sequence	2334 BP;	356 A;	803 C;	736 G;	439 T;
	Query Match	2.9%;	Score 58;	DB 37;	Length 2334;	
	Best Local Similarity	59.9%;	Pred. No. 6.84e-117;			
	Matches	175;	Conservative	0;	Mismatches	11; Indels 0; Gaps 0;
Df	439	tgtgccgcgcctgcgcttacacctaagcacaatgctgctgctggagatccaatgcgc	498			
Qy	401	TGTGCCGCAGCCTGCAATACCAAGAACATGGCGTGGCCCAACCTGCTGGGCCACGACCA	460			
Df	499	agacagagggcgctggaggtgacacagtctgtgctgctgggagatccaatgcgc	558			

RESULT	5	
ID	T76363	standard; DNA; 172 BP.
AC	AC	15-SEP-1997
DE	15-SEP-1997	(first entry)
DE	Human interleukin 8 antisense oligonucleotide.	
KW	Asthma; airway epithelium; adenosine free; cystic fibrosis;	
KW	chronic obstructive pulmonary disease; bronchitis; ss.	
OS	Synthetic.	
PN	W09640162-A1.	
PD	19-DEC-1996.	
PF	06-JUN-1996;	U09306.
PR	07-JUN-1995;	US-474497.
PA	(UYEC-) UNIV EAST CAROLINA.	
PI	Metzger WJ.;	Nyce JW;
DR	WPI: 97-051871/05.	
PT	Treatment of airway diseases such as asthma - by topically applying	
PT	adenosine-free antisense oligo:nucleotide to airway epithelium of	
PT	subject	
PS	Claim 5; page 36: 71pp: English.	
CC	A method for treating airway disease in a subject has been produced,	
CC	which involves the topical administration of an essentially adenosine	
CC	free antisense oligonucleotide (ON) to the airway epithelium of the	
CC	subject. The present sequence is an antisense oligonucleotide specific	
CC	for the human interleukin 8, targeted at the initiation codon. The	
CC	method can be used to treat airway diseases such as cystic fibrosis,	
CC	asthma, chronic obstructive pulmonary disease, bronchitis and other	
CC	airway diseases characterised by an inflammatory response. By	
CC	eliminating adenosine from the antisense ON, its liberation upon	
CC	antisense degradation is prevented, thereby preventing adenosine-	
CC	induced bronchoconstriction in patients with hyper-reactive airways.	
CC	Sequence 172 BP: 0 A; 35 C; 42 G; 39 T;	

RESULT 7
ID Q51746 standard; cDNA; 91 BP.
AC Q51746;
DT 31-MAY-1994 (first entry)

DE	Oligonucleotide probe MK14-A	
KW	Oligonucleotide: DNA probe; mycobacteria; disease diagnosis;	
KW	ss.	
OS	Synthetic.	
PN	EP-571911-A.	
PD	01-DEC-1993.	
PF	24-MAY-1993; 108325.	
PR	26-MAY-1992; US-889651.	
PA	(BECT) BECTON DICKINSON CO.	
PI	Shank DD, Spears PA;	
PT	WPI; 93-378844/48.	
PT	New oligo:nucleotide probes specific for Mycobacteria - used	
PT	detection and amplification of Mycobacteria nucleic acid in	
PT	samples	
PS	Claim 3; Page 14; 23pp; English.	
CC	Oligonucleotide probe MK14-A consists of nucleotides 5-95 of	
CC	(051735). It hybridized to all spp. of mycobacteria tested,	
CC	cross reacted to a few non-mycobacterial spp. The probe may	
CC	be useful as an initial screen for mycobacterial infection.	
CC	See also Q51735-45 and Q51747-59.	
SQ	Sequence 91 BP: 5 A: 17 C: 15 G: 4 T:	

```

Query Match      2.08; Score 41; DB 9; Length 91;
Best Local Similarity 0.08; Pred. No. 2.88e-07;
Matches 0; Conservative 44; Mismatches 3; Indels 0; Gaps 0;

Db      12 svhsyvvvvhvshhsvhhvhhvsvvvvhhvhhvhhvhhvhhv 58
Cp      1093 GCCCTTCTGCCCGCTTCACGAGGTGATCACCAGCTCCCAACC 1047

```

RESULT	8	
ID	T76186	standard; DNA; 100 BP.
AC	T76186;	
DT	12-SEP-1997	(first entry)
DE	Human IL4 receptor antisense oligonucleotide.	
KW	Asthma; airway epithelium; adenosine free; cystic fibrosis;	
KW	chronic obstructive pulmonary disease; bronchitis; interleukin; ss.	
OS	Synthetic.	
PN	WO9640162-A1.	
PD	19-DEC-1996.	
PF	06-JUN-1996;	UO9306.
PR	07-JUN-1995;	US-474497.
PA	(UYEC-) UNIV EAST CAROLINA.	
PI	Metzger WJ, Nyce JW;	
DR	MP1: 97-051871/05.	
PT	Treatment of airway diseases such as asthma - by topically applying	
PT	adenosine-free antisense oligo:nucleotide to airway epithelium of	
PT	subject	
PS	Example 5: Page 29; 71pp: English.	
CC	A method for treating airway disease in a subject has been produced,	
CC	which involves the topical administration of an essentially adenosine	
CC	free antisense oligonucleotide (ON) to the airway epithelium of the	
CC	subject. The present sequence is an antisense oligonucleotide specific	
CC	for the human IL4 receptor, targeted at the initiation codon. The method	
CC	can be used to treat airway diseases such as cystic fibrosis, asthma,	
CC	chronic obstructive pulmonary disease, bronchitis and other airway	
CC	diseases characterised by an inflammatory response. By eliminating	
CC	adenosine from the antisense ON, its liberation upon antisense	
CC	degradation is prevented, thereby preventing adenosine-induced	
CC	bronchoconstriction in patients with hyper-reactive airways.	
CC	Sequence 100 BP: 0 A: 33 C: 30 G: 10 T:	

	Query Match	2.0%	Score 40;	DB 32;	Length 100;
	Best Local Similarity	43.3%;	Pred. No. 8,76e-07;		
	Matches	39;	Conservative	26;	Mismatches 25; Indels 0; Gaps 0;
Db	7	bgbgcttcgcctcgcgcggcgcbgcggtccbcggagbbcbgagbccbgagcbbg	66	:	: :::
Qy	61	CGCGGGGCTTCAGCGCCCTTCGCCGCGTCTCCCGGTTC	120	:	: :::
Db	67	cbbccecbttgggbtgcbgcggcbbc	96	:	: :::

QY 121 GCCGCCGGCTGCCAGCTTTTCGGGGCCCCG 150

```

RESULT          9
ID      Q70466 standard; DNA; 114 BP.
AC      Q70466;
DT      05-APR-1995 (first entry)
DE      Generic DNA sequence to generate a random TSAR-9 petide library.
KW      TSAR; totally synthetic affinity reagent; synthetic; binding domain;
KW      effector domain; concatenated heterofunctional protein; linker;
KW      direct; rapid; detection; screening; treatment; generic; ss.
OS      Synthetic.
FH      Key
FT      Location/Qualifiers
FT      misc_feature      55..60
FT      /*tag= a
FT      /note= "this sequence represents 'Z'; Z can be a
FT      sequence of 6, 9 or 12 nucleotides (see
FT      comments)"
FT      WO9418318-A.
PN      18-AUG-1994.
PD      PD
PF      01-FEB-1994; U00977.
PR      01-FEB-1993; US-013416.
PR      30-DEC-1993; US-176500.
PR      31-JAN-1994; US-189331.
PA      (UYNC-) UNIV NORTH CAROLINA.
PI      Fowlkes DM, Kay BK;
DR      WPI; 94-279739/34.

```

identifying proteins or peptide(s) which bind a ligand - by
screening a recombinant vector library expressing fusion proteins
PT comprising a binding domain and an effector domain
PS Disclosure: Page 35; 255pp; English.
CC
CC 0704666 is a generic DNA sequence used to generate random TSAR (Totally
CC Synthetic Affinity Reagents) peptides. This generic formula can also be
CC represented as follows: X(NNB)10(TGC)(NNB)10(TGC)2(NNB)42(NNB)38(TGC)(NNB)
CC -9Y. X and Y are flanking restriction sites (X is not the same as Y)
CC that are not specified further. Other generic sequences are shown in
CC 0704666-68. Other specific peptides generated by these generic sequences
CC are shown in #65151-54. TSARs are concatenated heterofunctional proteins
CC or peptides, comprising at least two functional regions - a binding
CC domain with affinity for a ligand and a second effector peptide portion
CC that is chemically or biologically active. They may further comprise a
CC linker peptide between the 2 domains. The oligonucleotides are also
CC designed so that the expressed peptide contains 2 or 4 cysteine residues
CC positioned in, or flanking, the unpredicted or variant residues. These
CC residues confer some degree of conformational rigidity to the peptides.
CC The TSARs or compens. comprising a TSAR binding domain can be used in
CC vivo to deliver a chemically or biologically active moiety, eg. metal
CC ion, radioisotope, peptide, toxin or enzyme, to the specific target or
CC on the cell. They can also replace the function of macromolecules, eg.
CC monoclonal or polyclonal antibodies and therefore circumvent the need for
CC complex methods of hybridoma formation or in vivo antibody production.
CC The TSARs are easily characterised and have designed activity allowing
CC direct and rapid detection in a screening process.
SQ Sequence 114 BP; 0 A; 4 C; 4 G; 4 T.

Query Match 2.0%; Score 40; DB 12; Length 114;
Best Local Similarity 9.1%; Pred. No. 8.76e-07;
Matches 10; Conservative 31; Mismatches 69; Indels 0; Gaps 0;

[illegible]

RESULT 10
ID Q70468 standard; DNA; 114 BP.
AC Q70468;
DT 05-APR-1995 (first entry)


```
DE Generic DNA sequence to generate a random TSAR peptide library.
KW TSAR; totally synthetic affinity reagent; synthetic; binding domain;
KW effector domain; concatenated heterofunctional protein; linker;
KW direct; rapid; detection; screening; treatment; generic; ss.
OS Synthetic.
FH Key Location/Qualifiers
FT misc_feature 55..60
FT /*tag= a
FT /*note= "this sequence represents 'z'; z can be a
FT sequence of 6, 9 or 12 nucleotides (see
FT comments)"
PN WO9418318-A.
PD 18-AUG-1994. U00977.
PF 01-FEB-1994; US-013416.
PR 01-FEB-1993; US-013416.
PR 30-DEC-1993; US-176500.
PR 31-JAN-1994; US-189331.
PA (UUNC-) UNIV NORTH CAROLINA.
PI Fowlkes DM, Kay BK;
WPI: 94-279739/34.
DR P-PSDB: R65154.
DR Identifying proteins or peptide(s) which bind a ligand - by
PT screening a recombinant vector library expressing fusion proteins
PT comprising a binding domain and an effector domain
PS Disclosure: Page 35; 255pp; English.
CC Q70468 is a generic DNA sequence used to generate random TSAR (Totally
CC Synthetic Affinity Reagents) peptides. This generic formula can also be
CC represented as follows: X(NNB)11(TGC)(NNB)82(NNB)10V. X
CC and Y are flanking restriction sites (X is not the same as Y) that are
CC not specified further. Other generic sequences are shown in Q70466-68.
CC Other specific peptides generated by these generic sequences are shown in
CC Q70465-64. TSARs are concatenated heterofunctional proteins or peptides,
CC comprising at least two functional regions - a binding domain with
CC affinity for a ligand and a second effector peptide portion that is
CC chemically or biologically active. They may further comprise a linker
CC peptide between the 2 domains. The oligonucleotides are also designed so
CC that the expressed peptide contains 2 or 4 cysteine residues positioned
CC in, or flanking, the unpredicted or variant residues. These residues
CC confer some degree of conformational rigidity to the peptides. The TSARs
CC or compans. comprising a TSAR binding domain can be used in vivo to
CC deliver a chemically or biologically active moiety, eg. metal ion,
CC radioisotope, peptide, toxin or enzyme, to the specific target or on the
CC cell. They can also replace the function of macromolecules, eg.
CC monoclonal or polyclonal antibodies and therefore circumvent the need
CC for complex methods of hybridoma formation or in vivo antibody
CC production. The TSARs are easily characterised and have designed activity
CC allowing direct and rapid detection in a screening process.
SQ Sequence 114 BP; 0 A; 2 C; 2 G; 2 T;

Query Match 2.0%; Score 40; DB 12; Length 114;
Best Local Similarity 5.4%; Pred. No. 8.76e-07;
Matches 6; Conservative 34; Mismatches 72; Indels 0; Gaps 0;

Db 3 bnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnnn 62
QY 199 CCGCTCGCCCTTCGCGGCTCGCTCTCCCTCTCGCCCTCGCGGCTCGCCGACGATG 258
Db 63 bnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnb 114
QY 259 CTGAGGGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCC 310

RESULT 11
ID N81164 standard; DNA; 204 BP.
AC N81164;
DT 08-NOV-1990 (first entry)
DE Base substituted E.coli beta-galactosidase alpha-fragment.
KW E.coli beta galactosidase alpha-fragment; base substitutions; ss.
OS Escherichia coli.
FH Key Location/Qualifiers
FT misc_feature 19..69
FT /*tag= a
FT /function= multiple cloning site
```

```
FT primer_bind 187..204
FT /*tag= b
FT EP-285123-A.
PD 05-MAY-1988.
PF 30-MAR-1988; 105163.
PR 03-APR-1987; US-034819.
PA (SUSO) SUOMEN SOKERI OY.
PI Lehtovaara P, Knowles J, Koivula A, Bamford J, Reinikainen T;
WPI: 88-279927/40.
DR Introducing random point mutations into nucleic acids -
PT by prep of single stranded template, annealing a primer, elongation,
PT misincorporation, completion of molecules and screening.
PS Disclosure: P; English.
CC Random point mutations were introduced into the alpha fragment of
CC E.coli beta-galactosidase. The wild type sequence was obtained as a
CC single stranded template and an oligonucleotide was hybridised to
CC it to generate a popn of DNA molecules which terminate at all
CC possible nucleotide positions within a specified region. The
CC variable 3' ends generated in this way are used as primers for
CC reverse transcriptase. Nucleotides are misincorporated by the
CC transcriptase and the molecules are completed to forms that can be
CC amplified and then expressed in a suitable host-vector system.
CC The sequence covers all 176 diff base substitutions, most of which
CC occurred singularly in any given mutant.
CC See also F80375.
SQ Sequence 204 BP; 21 A; 47 C; 17 G; 11 T; 108 Others;

Query Match 2.0%; Score 40; DB 1; Length 204;
Best Local Similarity 15.6%; Pred. No. 8.76e-07;
Matches 20; Conservative 57; Mismatches 51; Indels 0; Gaps 0;

Db 59 ggyywcgagcygcaaycdhvcgcygrrtthhyrrmbnvyrdynrsdaaayccyr 118
Cp 343 GCCCAAGAGGAAGAGCGCGCGCCGAGCCGAGCCGAGCGAGCGAGGAGGAGCA 284
Db 119 rsvkydcynachdhvbybbvynvnhnncnccbnhvhvbnhnnhrnwayvrhd 178
Cp 283 GCAGGAGGAGGAGCGCGCGCCCTGCGATCTGTCGCGCGCGGCGGCGGAGGAGGA 224
Db 179 arddgvhc 186
Cp 223 GCGGAGCC 216

RESULT 12
ID Q70469 standard; DNA; 114 BP.
AC Q70469;
DT 07-APR-1995 (first entry)
DE Generic DNA sequence to generate a random TSAR peptide library.
KW TSAR; totally synthetic affinity reagent; synthetic; binding domain;
KW effector domain; concatenated heterofunctional protein; linker;
KW direct; rapid; detection; screening; treatment; generic; ss.
OS Synthetic.
FH Key Location/Qualifiers
FT misc_feature 55..60
FT /*tag= a
FT /*note= "this sequence represents 'z'; z can be a
FT sequence of 6,9 or 12 nucleotides (see
FT comments)"
PN WO9418318-A.
PD 18-AUG-1994. U00977.
PF 01-FEB-1994; US-013416.
PR 01-FEB-1993; US-013416.
PR 30-DEC-1993; US-176500.
PR 31-JAN-1994; US-189331.
PA (UUNC-) UNIV NORTH CAROLINA.
PI Fowlkes DM, Kay BK;
WPI: 94-279739/34.
DR Identifying proteins or peptide(s) which bind a ligand - by
PT screening a recombinant vector library expressing fusion proteins
PT comprising a binding domain and an effector domain
PS Disclosure: Page 35; 255pp; English.
CC Q70469 is a generic DNA sequence used to generate random TSAR peptide
```

comprising at least two functional regions, a binding domain with affinity for a ligand and a second effector peptide portion that is chemically or biologically active. They may further comprise a linker peptide between the 2 domains. The oligonucleotides are also designed so that the expressed peptide contains 2 or 4 cysteine residues positioned in, or flanking, the unpredicted or variant residues. These residues confer some degree of conformational rigidity to the peptides. The TSARs

The oligonucleotides are also designed so that the expressed peptide contains 2 or 4 cysteine residues positioned in, or flanking, the unpredicted or variant residues. These residues confer some degree of conformational rigidity to the peptides. The TSARs or compns. comprising a TSAR binding domain can be used *in vivo* to deliver a chemically or biologically active moiety, eg. metal ion, radioisotope, peptide, toxin or enzyme, to the specific target or on the cell. They can also replace the function of macromolecules, eg. monoclonal or polyclonal antibodies and therefore circumvent the need for complex methods of hybridoma formation or *in vivo* antibody production. The TSARs are easily characterised and have designed activity allowing direct and rapid

THIS PAGE BLANK (USPTO)

WATER

(TW)

Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd
MPsrch_nn n.a. - n.a. database search, using Smith-Waterman algorithm
Run on: Wed Aug 5 06:20:04 1998; MasPar time 2338.81 Seconds
Tabular output not generated.
Title: >US-08-848-439-1
Description: (1-2027) from US08848439.seq
Perfect Score: 2027
N.A. Sequence: 1 GAATTCGGCCTTCATGGCT.....AAAAAAGGCGCGC 2027
Comp: CTTAAGCGGAGTACCGGA.....TTTTTTTTTCGCCGCGC

Scoring table: TABLE default
Gap 6
Nmatch STD : Dbase 0; Query 0
Searched: 1832099 segs, 700269816 bases x 2
Post-processing: Minimum Match 0%
Listing first 45 summaries
Database: emb1-est55
1:em_est10 2:em_est11
genbank-est107
Database: 3:gb_est1 4:gb_est10 5:gb_est11 6:gb_est12 7:gb_est13
8:gb_est14 9:gb_est15 10:gb_est16 11:gb_est17
12:gb_est18 13:gb_est19 14:gb_est2 15:gb_est20
16:gb_est21 17:gb_est22 18:gb_est23 19:gb_est3
20:gb_est4 21:gb_est5 22:gb_est6 23:gb_est7 24:gb_est8
25:gb_est9 26:gb_gss

Statistics: Mean 12.269; Variance 3.137; scale 3.912
Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
C 1	553	27.3	579	7	AA528226 nh92a05.s1 NCI_CGAP_Br	0.00e+00
C 2	479	23.6	505	24	ze88f06.s1 Soares feta	0.00e+00
C 3	473	23.3	480	21	Yy64d09.r1 Homo sapien	0.00e+00
C 4	472	23.3	505	23	ze03g10.r1 Soares feta	0.00e+00
C 5	468	23.1	473	8	nh88g03.s1 NCI_CGAP_Br	0.00e+00
C 6	468	23.1	474	15	of62a09.s1 NCI_CGAP-Co	0.00e+00
C 7	453	22.3	460	23	zd70e11.s1 Soares feta	0.00e+00
C 8	434	21.4	443	7	zx08c06.s1 Soares tota	0.00e+00
C 9	433	21.4	482	18	ch10h11.s1 NCI_CGAP-Co	0.00e+00
C 10	425	21.0	427	5	zw51g08.s1 Soares tota	0.00e+00
C 11	420	20.7	518	17	on47d08.s1 NCI_CGAP-Co	0.00e+00
C 12	416	20.5	446	21	Yy64d09.s1 Homo sapien	0.00e+00
C 13	413	20.4	462	23	zd70e11.r1 Soares feta	0.00e+00

C 14	409	20.2	409	7	AA449749	zx7e10.s1 Soares tota	0.00e+00
C 15	408	20.1	445	20	N32415	Yw83c06.s1 Homo sapien	0.00e+00
C 16	407	20.0	562	7	AA4490372	zx07e10.r1 Soares tota	0.00e+00
C 17	405	20.0	405	8	AA505877	ni01a05.s1 NCI_CGAP_Br	0.00e+00
C 18	402	19.8	408	18	AA976403	Oq88f09.s1 NCI_CGAP_Ki	0.00e+00
C 19	388	19.1	388	7	AA528219	nj16h09.s1 NCI_CGAP_Pr	0.00e+00
C 20	366	18.1	422	5	AA431858	zw51g08.r1 Soares tota	0.00e+00
C 21	366	18.1	458	24	AA105749	ml84a03.r1 Stratagene	0.00e+00
C 22	362	17.9	378	10	AA429960	zw58g10.r1 Soares tota	0.00e+00
C 23	363	17.9	433	22	AA024772	ze76g10.s1 Soares feta	0.00e+00
C 24	359	17.7	426	20	H87071	Ys74d07.r1 Homo sapien	0.00e+00
C 25	350	17.3	359	4	AA295018	EST100474 Pancreas tum	0.00e+00
C 26	337	16.6	406	19	H16121	Y128c05.r1 Homo sapien	0.00e+00
C 27	331	16.3	340	6	HUM296B07B	Human aorta cDNA 5'-en	0.00e+00
C 28	329	16.2	396	25	AA073862	mm97f07.r1 Stratagene	0.00e+00
C 29	328	16.2	414	22	AA024771	ze76g10.r1 Soares feta	0.00e+00
C 30	323	15.9	329	12	AA705737	zf41b02.s1 Soares feta	0.00e+00
C 31	322	15.9	434	19	H44092	Y073a07.r1 Homo sapien	0.00e+00
C 32	321	15.8	321	4	AA335712	EST40158 Epididymus Ho	0.00e+00
C 33	320	15.8	328	4	AA295688	EST100887 Pancreas tum	0.00e+00
C 34	321	15.8	469	19	H15818	Y128c05.s1 Homo sapien	0.00e+00
C 35	319	15.7	326	20	N32424	Yw83e07.s1 Homo sapien	0.00e+00
C 36	316	15.6	330	6	HUM230F10B	Human aorta cDNA 5'-en	0.00e+00
C 37	313	15.4	344	21	N56835	Yw83c06.r1 Homo sapien	0.00e+00
C 38	299	14.8	301	6	HUM303B01B	Human aorta cDNA 5'-en	0.00e+00
C 39	301	14.8	301	10	AA602994	np31g01.s1 NCI_CGAP_Pr	0.00e+00
C 40	298	14.8	313	6	HUM304B12B	Human aorta cDNA 5'-en	0.00e+00
C 41	299	14.7	303	6	HUM347H05B	Human aorta cDNA 5'-en	0.00e+00
C 42	296	14.6	300	4	AA347786	EST54420 Fetal heart I	0.00e+00
C 43	294	14.5	300	6	HUM307G04B	Human aorta cDNA 5'-en	0.00e+00
C 44	292	14.4	302	4	AA330758	EST34493 Embryo, 6 wee	0.00e+00
C 45	290	14.3	293	6	HUM238C04B	Human aorta cDNA 5'-en	0.00e+00

ALIGNMENTS

RESULT 1
LOCUS AA528226 579 bp mRNA EST 05-AUG-1997
DEFINITION nh92a05.s1 NCI_CGAP_Br.1.1 Homo sapiens cDNA clone IMAGE:965936,
mRNA sequence.
ACCESSION AA528226
NID g2270295
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryote; mitochondrial eukaryotes; Metazoa; Chordata;
Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Homnidae;
Homo.
REFERENCE 1 (bases 1 to 579)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert_Strausberg@nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: M. Bento Soares, Ph.D.
DNA Sequencing by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.lnl.gov/bbrp/image/image.html
Insert Length: 1268 Std Error: 0.00
Seq primer: -40m13 fwd. ET from Amersham
High quality sequence stop: 206.
Location/Qualifiers
1. .579
/organism="Homo sapiens"
/note="Vector: pT73D-Pac (Pharmacia) with a modified

FEATURES
source

polylinker; 1st strand cDNAs were prepared from pooled bulk primary tumor tissues and then primed with a Not I-B150(pT) primer. Double-stranded cDNA was ligated to EcoRI and Not I-B150(pT) and digested with Not I and cloned in the Not I-B150(pT) sites of the modified pT7h3 vector. The library is not normalized. (The normalized vector library is not shown.) (The normalized library was constructed by Aceto-Snars and M. Fatima Bonaguidi. This library is NCICAP Br2. "Library was constructed by Aceto-Snars and M. Fatima Bonaguidi."

BASE COUNT	198 a	80 c	104 g	197 t	
mrna					
ORIGIN					
		<1. ->579			
Query Match		27.3%;	Score 553;	DB 7;	Length 579;
Best Local Similarity		99.1%;	Pred. No. 0.00e+00;		
Matches	558;	Conservative	0;	Mismatches	5;
				Indels	0;
				Gaps	
Db	17	TTTTTTTTTTTTTTTTTTTTTTAGGTAAACACAGAGATGTAAGATTTTATATACAGAATAATAATGTT	76		
Cp	2019	TTTTTTTTTTTTTTTTTTTTTTAGGTAAACACAGAGATGTAAGATTTTATATACAGAATAATAATGTT	1960		
Db	77	TATCTGAAATATTTACAGTGTGGTTAAAGCAATATTTTTACAACCTTTTAAAGGTAAACT	138		
Cp	1959	TATCTGAAATATTTACAGTGTGGTTAAAGCAATATTTTACAACCTTTTAAAGGTAAACT	1900		
Db	137	ACTATGTATATTACAGTAAAGTACAAATGGGTTTTAAATTTGCAAAAGTTAAGTAAGAAATG	196		
Cp	1899	ACTATGTATATTACAGTAAAGTACAAATGGGTTTTAAATTTGCAAAAGTTAAGTAAGAAATG	1840		
Db	197	TTTTAAACAAAGGCTTAAAGTACTCAAGTCAATATAAAATTTATATCTTTTGCCCTTTTAC	256		
Cp	1839	TTTTAAACAAAGGCTTAAAGTACTCAAGTCAATATAAAATTTATATCTTTTGCCCTTTTAC	1780		
Db	257	TTGAAGAAATCATGCPATAGAAATGGTTAAATGCTGTTCTAATAATGGAAGTATTGTAGC	316		
Cp	1779	TTGAAGAAATCATGCPATAGAAATGGTTAAATGCTGTTCTAATAATGGAAGTATTGTAGC	1720		
Db	317	TGGATGTGTATACATCTAACAGTTTAAAGTCCCATTTGAAGGTATAAAATGATGAATGTT	376		
Cp	1719	TGGATGTGTATACATCTAACAGTTTAAAGTCCCATTTGAAGGTATAAAATGATGAATGTT	1660		
Db	377	GTAAAGCTTTAGACACTGAGTCTCAGTCTGGAGCTGATGAAGATGTGAGATAACAGCCAG	436		
Cp	1659	GTAAGACTTTAGACACTGAGTCTCAGTCTGGAGCTGATGAAGATGTGAGATAACAGCCAG	1600		
Db	437	CTTTTATCTCAACAGGTTTGTGACCCACAAGTTTGGGCCACAGAGAAAATTTGAAGCCATT	496		
Cp	1599	CTTTTATCTCAACAGGTTTGTGACCCACAAGTTTGGGCCACAGAGAAAATTTGAAGCCATT	1540		
Db	497	TGCATGTTATGACAACCTCAGTGGGAAGTGAAAATCCGCTGACTCAAAACCCACAAAACAC	556		
Cp	1539	TGCATGTTATGACAACCTCAGTGGGAAGTGAAAATCCGCTGACTCAAAACCCACAAAACAC	1480		
Db	557	AACCAACGAGCCCAAGTCCCG	579		
Cp	1479	AACCAACGAGCCCAAGTCCCG	1457		

[illegible]

REFERENCE AUTHORS	TITLE JOURNAL COMMENT

vertebrates: Eutheria; Primates; Catarrhini; Hominoidea: Hominidae: Homo
1 (bases 1 to 595)
Hillier, L., Clark, M., Dubouche, T., Ellison, K., Hawkins, M.,
Holman, M., Hultman, M., Kucaba, T., Le, M., Lennon, G., Marra, M.,
Parsons, J., Rifkin, L., Rohlfing, T., Tan, F., Trevaskis, E.,
Waterson, R., Williamson, A., Wohldmann, P. and Wilson, R.
Washington: EST Project
WashU-Merck
Unpublished (1995)

Contact: Wilson RK
WashU-Merck EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810

This clone is available royalty-free through LLNL ; contact the
Email: estewartson.wustl.edu
IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 1171 Std Error: 0.00
Seq primer: -40M13 fwd. from Amersham
High quality sequence stop: 324.

FEATURES	SOURCE
----------	--------

1. .505
/organism="Homo sapiens"
/note="Organ: heart; Vector: pT73D (Pharmacia) with a
modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st
strand cDNA was primed with a Not I - oligo(dT) primer [5
TGTACCAACTTCGAAGTGGAGCGCGGATTTTTTTTTTTT 3'],
TGTACCAACTTCGAAGTGGAGCGCGGATTTTTTTTTTTT 3'],
double-stranded cDNA was size selected, ligated to Eco RI
adapters (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of a modified pT73 vector
(Pharmacia). Library went through one round of
normalization to a Cot = 5. Library constructed by
M.Fatima Bonaldo. This library was constructed from the
same tissue as the fetal lung library, Soares fetal lung
cDNA library."

mRNA	BASE COUNT	ORIGIN
179 a	61 c	100 g
162 t		
3 others		

Query Match	23.6%	Score	479;	DB	24;	Length	505;
Best Local Similarity	99.0%;	Pred. No.	0.00e+00;				
Matches	498;	Conservative	0;	Mismatches	2;	Indels	3;
Db	2	TAGGTAAACAGGATGTAAGT	TTATATACAAAGATATATATGTTTATCTCGAATATTAC	61			
Cp	2003	TAGTAAACAGGATGTAAGT	TTATATACAAAGATATATGTTTATCTCGAATATTAC	1944			
Db	62	AGTGTGGTAAAGCAATATTTTACAACTTTTAAAGGTAACACTACTATGTATATTACAG	121				
Cp	1943	AGTGTGGTAAAGCAATATTTTACAACTTTTAAAGGTAACACTACTATGTATATTACAG	1884				
Db	122	GTAAGCTACAAATGGGTTTAAATTTGCAAAAGTTAAGTAAGAAATGTTTTAAACAAGGCTTA	181				
Cp	1883	GTAAGCTACAAATGGGTTTAAATTTGCAAAAGTTAAGTAAGAAATGTTTTAAACAAGGCTTA	1824				
Db	182	AAGTACTCAAGTCAATTTAAAATTTATATCTTTTGGCTTTTACTTGAAGAAATCATGCT	241				
Cp	1823	AAGTACTCAAGTCAATTTAAAATTTATATCTTTTGGCTTTTACTTGAAGAAATCATGCT	1764				
Db	242	ATAGAAATGGTTAATGTGCTTCTTAATAAATGGAAGTATTCTAGCTGGAATGTGATACATG	301				
Cp	1763	ATAGAAATGGTTAATGTGCTTCTTAATAAATGGAAGTATTCTAGCTGGAATGTGATACATG	1704				
Db	302	TAAACAGTTTAAGTTCCCAATGCAAGGTATAAAATGATGAATTTGTTGAAGACTTAGACACT	361				
Cp	1703	TAAACAGTTTAAGTTCCCAATGCAAGGTATAAAATGATGAATTTGTTGAAGACTTAGACACT	1644				

normalization to a Cot = 5. Library constructed by M.Fatima Bonaldo. This library was constructed from the same fetus as the fetal lung library, Soares fetal lung NBHL19W." /db_xref="taxon:9606" /clone="357954" /clone_lib="Soares fetal heart NBHL19W" /sex="unknown" /dev_stage="19 weeks" /lab_host="DH10B (ampicillin resistant)" <1..>305

mRNA 149 a 95 c 81 g 180 t
BASE COUNT
ORIGIN

Query Match 23.3%; Score 472; DB 23; Length 505;
Best Local Similarity 98.8%; Pred. No. 0.00e+00;
Matches 498; Conservative 0; Mismatches 2; Indels 4; Gaps 4;
Db 1 ATTTTATGAAGTTTAAATAGCTACCTTTAAAGCTAGTTTGAATAGTGCAACTGTG 60
QY 1402 ATTTTATGAAGTTTAAATAGCTACCTTTAAAGCTAGTTTGAATAGTGCAACTGTG 1461
Db 61 ACTTGGGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 120
QY 1462 ACTTGGGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1521
Db 121 AGTTTGTCTAATCAATGCAAAATGCTTCAATTTCTCTGTGCCCCAAACTTGTGGGTCA 180
QY 1522 AGTTTGTCTAATCAATGCAAAATGCTTCAATTTCTCTGTGCCCCAAACTTGTGGGTCA 1581
Db 181 AACCTGTGTGAGATAAGCTGGCTGTATCTCAACATCTTCATCAGCTCCAGACTGAGAC 240
QY 1582 AACCTGTGTGAGATAAGCTGGCTGTATCTCAACATCTTCATCAGCTCCAGACTGAGAC 1641
Db 241 TCAGTGTCTAAGTCTTACAAACATTCATCATTTTATACCTTCAATGGGAACCTTAACTGT 300
QY 1642 TCAGTGTCTAAGTCTTACAAACATTCATCATTTTATACCTTCAATGGGAACCTTAACTGT 1701
Db 301 TACATGTATCATCTCCAGCTACAACTACTTCCATTTTATTAAGAGCACATTAACCATTT 360
QY 1702 TACATGTATCATCTCCAGCTACAA-TACTTCCATTTATTA-GAAGCACATTAACCATTT 1759
Db 361 CTATAGCATGATTTCTCAAGTAAAGGCAAGATATAAATTTATTAATGACTTGAGT 420
QY 1760 CTATAGCATGATTTCTCAAGTAAAGGCAAGATATAAATTTATTAATGACTTGAGT 1819
Db 421 ACTTTAAGCCTTGTATAACATTTCTTACTTAACCTTTTGCAATTAACCCATTTGCTAG 480
QY 1820 ACTTTAAGCCTTGTATAACATTTCTTACTTAACCTTTTGCAATTAACCCATTTGCTAG 1878
Db 481 CTTACCCGGTAATATACATAGTAG 504
QY 1879 CTTACCTG-TAATATACATAGTAG 1901

RESULT 5
LOCUS AA513750 473 bp mRNA EST 20-AUG-1997
DEFINITION nh88g03.s1 NCI_CGAP_Brl.1 Homo sapiens cdna clone IMAGE:965620,
mRNA sequence.
ACCESSION AA513750
NID 92252162
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryotes; mitochondrial eukaryotes; Metazoa; Chordata;
Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominidae;
Homo.
REFERENCE 1 (bases 1 to 473)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT

Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert.Strausberg@nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
Emmert-Buck, M.D., Ph.D.
CDNA Library Preparation: M. Bento Soares, Ph.D.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html

Insert Length: 2969 Std Error: 0.00
Seq primer: -40m13 fwd. Et from Amersham
High quality sequence stop: 472.

FEATURES
source

Location/Qualifiers
1..473
/organism="Homo sapiens"
/note="Vector: pT7T3D-Pac (Pharmacia) with a modified
polylinker; 1st strand cDNA was prepared from pooled bulk
breast tumor tissue, and was then primed with a Not I -
oligo(dT) primer. Double-stranded cDNA was ligated to Eco
RI adaptors (Pharmacia), digested with Not I and cloned
into the Not I and Eco RI sites of the modified pT7T3
vector. Library is not normalized. (The normalized
version of this library is NCI_CGAP_Br2.) Library was
constructed by Bento Soares and M. Fatima Bonaldo."
/db_xref="taxon:9606"
/clone="IMAGE:965620"
/clone_lib="NCI_CGAP_Brl.1"
/sex="female, pooled"
/tissue_type="breast"
/lab_host="DH10B"
mRNA 169 a 57 c 91 g 156 t
BASE COUNT
ORIGIN

Query Match 23.1%; Score 468; DB 8; Length 473;
Best Local Similarity 99.6%; Pred. No. 0.00e+00;

Matches 470; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Db 1 AGGTAAACAGGATGTAAGTTTATATACAGAAATATAATGTTTATCTGAAATATTACA 60
Cp 2002 AGGTAAACAGGATGTAAGTTTATATACAGAAATATAATGTTTATCTGAAATATTACA 1943
Db 61 GTGTGGTTAAAGCAATATTTTACAACTTTTAAAGTAACTACTATGTATATTACAGG 120
Cp 1942 GTGTGGTTAAAGCAATATTTTACAACTTTTAAAGTAACTACTATGTATATTACAGG 1883
Db 121 TAAGCTACATGGGTTTAAATTTGCAAAAGTTAAGTAAAGAAATGTTTAAACAAGGCTTAA 180
Cp 1882 TAAGCTACATGGGTTTAAATTTGCAAAAGTTAAGTAAAGAAATGTTTAAACAAGGCTTAA 1823
Db 181 AGTACTCAAGTCAATATAAAATTTATATCTTTTGCTTTTACTTGAAGAAATCATGCTA 240
Cp 1822 AGTACTCAAGTCAATATAAAATTTATATCTTTTGCTTTTACTTGAAGAAATCATGCTA 1763
Db 241 TAGAAATGGTTAATGCTTCTTAATAAATGGAAGTATTGTAGCTGGAATGTGATACATGT 300
Cp 1762 TAGAAATGGTTAATGCTTCTTAATAAATGGAAGTATTGTAGCTGGAATGTGATACATGT 1703
Db 301 AACAGTTTAAAGTTCCCATGGAAGGTATAAAATGATGAATTTGTTGAAGACTTAGACACTG 360
Cp 1702 AACAGTTTAAAGTTCCCATGGAAGGTATAAAATGATGAATTTGTTGAAGACTTAGACACTG 1643
Db 361 AGTCTCGCTCGAGCTGATGAAGATTTGAGATACAGCCAGCTTTTATCTCAACAGGTT 420
Cp 1642 AGTCTCGCTCGAGCTGATGAAGATTTGAGATACAGCCAGCTTTTATCTCAACAGGTT 1583
Db 421 TTGTCACCCACAAAGTTGGGCCCCAGAGAAAATTGAAGCAAAATTTGCATGTTA 472
Cp 1582 TTGTCACCCACAAAGTTGGGCCCCAGAGAAAATTGAAGCAAAATTTGCATGTTA 1531

RESULT	6	AA857092	474 bp	mRNA	EST	09-MAR-1998				
LOCUS		of62a09.s1	NCI_CGAP_Co8	Homo sapiens	cdna clone	IMAGE:1434904 3',				
DEFINITION		mRNA sequence.								
ACCESSION		AA857092								
NID		92945394								
KEYWORDS		EST.								
SOURCE		human.								
ORGANISM		Homo sapiens								
		Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;								
		Primates; Catarrhini; Homnidae; Homo.								
REFERENCE		1 (bases 1 to 474)								
AUTHORS		NCI-CGAP	http://www.ncbi.nlm.nih.gov/ncicgap.							
TITLE		National Cancer Institute, Cancer Genome Anatomy Project (CGAP),								
		Tumor Gene Index								
JOURNAL		Unpublished (1997)								
COMMENT		Contact: Robert Strausberg, Ph.D.								
		Tel: (301) 496-1550								
		Email: Robert.Strausberg@nih.gov								
		Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.								
		Emmert-Buck, M.D., Ph.D.								
		CDNA Library Preparation: M. Bento Soares, Ph.D.								
		DNA Library Arrayed by: Greg Lennon, Ph.D.								
		Sequencing by: Washington University Genome Sequencing Center								
		Clone distribution: NCI-CGAP clone distribution information can be								
		found through the I.M.A.G.E. Consortium/LLNL at:								
		www-bio.llnl.gov/bbrp/image/image.html								
Seq primer:		-40ml3 fwd. ET from Amersham								
High quality sequence stop:		464.								
FEATURES		Location/Qualifiers								
source		1..474								
		/organism="Homo sapiens"								
		/note="Organ: colon; Vector: pT7T3D-Pac (Pharmacia) with a								
		modified polylinker; 1st strand cDNA was prepared from								
		colon adenocarcinoma, and was then primed with a Not I -								
		oligo(dT) primer. Double-stranded cDNA was ligated to Eco								
		RI adaptors (Pharmacia), digested with Not I and cloned								
		into the Not I and Eco RI sites of the modified pT7T3								
		vector. Library is normalized. Library was constructed by								
		Bento Soares and M. Fatima Bonaldo."								
		/db_xref="taxon:9606"								
		/clone="IMAGE:1434904"								
		/clone_lib="NCI_CGAP_Co8"								
		/tissue_type="adenocarcinoma"								
		/lab_host="DH10B"								
BASE COUNT		171 a	54 c	92 g	157 t					
ORIGIN										
Query Match		23.1%; Score 468; DB 15; Length 474;								
Best Local Similarity		99.8%; Pred.No. 0.00e+00;								
Matches		474; Conservative 0; Mismatches 0; Indels 1; Gaps 1;								
Db	1	AGGTAAACAGGATGTAAGTGTATATACAGAATAATAATGTTTCTCGAAATATTACA 60								
Cp	2002	AGGTAAACAGGATGTAAGTGTATATACAGAATAATAATGTTTCTCGAAATATTACA 1943								
Db	61	GTGTGGTTAAAGCAATATTTTACAACCTTTTAAAGGTAACTACTATGTATATTACAGG 120								
Cp	1942	GTGTGGTTAAAGCAATATTTTACAACCTTTTAAAGGTAACTACTATGTATATTACAGG 1883								

```

Cp 1762 TAGAATGGTTAATGCGCTCTTAATAAATGGAAGTATTGTTAGCTGGAATGTTGATACATGT 1703
Db 301 AACAGTTTAAAGTTCCCATTTAAAGGTATAAAATGATGAATTTGTTGAAGACTTTAGACACTG 360
|||
Cp 1702 AACAGTTTAAAGTTCCCATTTAAAGGTATAAAATGATGAATTTGTTGAAGACTTTAGACACTG 1643
|||
Db 361 AGTCTCAGTCTGGAGCTGATGAAGATGTTTGAGATTAACAGCCAGCTTTATCTCAACAGGTT 420
|||
Cp 1642 AGTCTCAGTCTGGAGCTGATGAAGATGTTTGAGATTAACAGCCAGCTTTATCTCAACAGGTT 1583
|||
Db 421 TTGTGACC-ACAAAGTTTGGCCACAGAGAAAAATTGAAGCAAAATTTGCATGTTTATGA 474
|||
Cp 1582 TTGTGACCCACAAGTTTGGCCACAGAGAAAATTTGAAGCAATTTGTCATGTTATGA 1528
|||

RESULT 7
LOCUS W72225 460 bp mRNA EST 17-OCT-1996
DEFINITION zd70e11.s1 Soares fetal heart NBHH19W Homo sapiens cDNA clone
346028 3', mRNA sequence.
ACCESSION W72225
NID 91382674
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryote; mitochondrial eukaryotes; Metazoa; Chordata;
Vertebrata; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 460)
AUTHORS Hillier,L., Clark,N., Dubucq,T., Elliston,K., Hawkins,M.,
Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,
Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F.,
Travaskis,E., Waterston,R., Williamson,A., Wohlmann,P. and
Wilson,R.
TITLE The WashU-Merck EST Project
JOURNAL Unpublished (1995)
COMMENT
Contact: Wilson RK
WashU-Merck EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available royalty-free through LLNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 1550 Std Error: 0.00
Seq primer: mob.REGA+ET
High quality sequence stop: 384.
Location/Qualifiers
1. .460
/organism="Homo sapiens"
/note="Organ: heart; Vector: pT73D (Pharmacia) with a
modified polylinker; Site_1: Not I; Site_2: Eco RI; Site_3:
Not I; Site_4: Not I; Site_5: Not I; Site_6: Not I; Site_7:
Not I; Site_8: Not I; Site_9: Not I; Site_10: Not I; Site_11:
Not I; Site_12: Not I; Site_13: Not I; Site_14: Not I; Site_15:
Not I; Site_16: Not I; Site_17: Not I; Site_18: Not I; Site_19:
Not I; Site_20: Not I; Site_21: Not I; Site_22: Not I; Site_23:
Not I; Site_24: Not I; Site_25: Not I; Site_26: Not I; Site_27:
Not I; Site_28: Not I; Site_29: Not I; Site_30: Not I; Site_31:
Not I; Site_32: Not I; Site_33: Not I; Site_34: Not I; Site_35:
Not I; Site_36: Not I; Site_37: Not I; Site_38: Not I; Site_39:
Not I; Site_40: Not I; Site_41: Not I; Site_42: Not I; Site_43:
Not I; Site_44: Not I; Site_45: Not I; Site_46: Not I; Site_47:
Not I; Site_48: Not I; Site_49: Not I; Site_50: Not I; Site_51:
Not I; Site_52: Not I; Site_53: Not I; Site_54: Not I; Site_55:
Not I; Site_56: Not I; Site_57: Not I; Site_58: Not I; Site_59:
Not I; Site_60: Not I; Site_61: Not I; Site_62: Not I; Site_63:
Not I; Site_64: Not I; Site_65: Not I; Site_66: Not I; Site_67:
Not I; Site_68: Not I; Site_69: Not I; Site_70: Not I; Site_71:
Not I; Site_72: Not I; Site_73: Not I; Site_74: Not I; Site_75:
Not I; Site_76: Not I; Site_77: Not I; Site_78: Not I; Site_79:
Not I; Site_80: Not I; Site_81: Not I; Site_82: Not I; Site_83:
Not I; Site_84: Not I; Site_85: Not I; Site_86: Not I; Site_87:
Not I; Site_88: Not I; Site_89: Not I; Site_90: Not I; Site_91:
Not I; Site_92: Not I; Site_93: Not I; Site_94: Not I; Site_95:
Not I; Site_96: Not I; Site_97: Not I; Site_98: Not I; Site_99:
Not I; Site_100: Not I; Site_101: Not I; Site_102: Not I; Site_103:
Not I; Site_104: Not I; Site_105: Not I; Site_106: Not I; Site_107:
Not I; Site_108: Not I; Site_109: Not I; Site_110: Not I; Site_111:
Not I; Site_112: Not I; Site_113: Not I; Site_114: Not I; Site_115:
Not I; Site_116: Not I; Site_117: Not I; Site_118: Not I; Site_119:
Not I; Site_120: Not I; Site_121: Not I; Site_122: Not I; Site_123:
Not I; Site_124: Not I; Site_125: Not I; Site_126: Not I; Site_127:
Not I; Site_128: Not I; Site_129: Not I; Site_130: Not I; Site_131:
Not I; Site_132: Not I; Site_133: Not I; Site_134: Not I; Site_135:
Not I; Site_136: Not I; Site_137: Not I; Site_138: Not I; Site_139:
Not I; Site_140: Not I; Site_141: Not I; Site_142: Not I; Site_143:
Not I; Site_144: Not I; Site_145: Not I; Site_146: Not I; Site_147:
Not I; Site_148: Not I; Site_149: Not I; Site_150: Not I; Site_151:
Not I; Site_152: Not I; Site_153: Not I; Site_154: Not I; Site_155:
Not I; Site_156: Not I; Site_157: Not I; Site_158: Not I; Site_159:
Not I; Site_160: Not I; Site_161: Not I; Site_162: Not I; Site_163:
Not I; Site_164: Not I; Site_165: Not I; Site_166: Not I; Site_167:
Not I; Site_168: Not I; Site_169: Not I; Site_170: Not I; Site_171:
Not I; Site_172: Not I; Site_173: Not I; Site_174: Not I; Site_175:
Not I; Site_176: Not I; Site_177: Not I; Site_178: Not I; Site_179:
Not I; Site_180: Not I; Site_181: Not I; Site_182: Not I; Site_183:
Not I; Site_184: Not I; Site_185: Not I; Site_186: Not I; Site_187:
Not I; Site_188: Not I; Site_189: Not I; Site_190: Not I; Site_191:
Not I; Site_192: Not I; Site_193: Not I; Site_194: Not I; Site_195:
Not I; Site_196: Not I; Site_197: Not I; Site_198: Not I; Site_199:
Not I; Site_200: Not I; Site_201: Not I; Site_202: Not I; Site_203:
Not I; Site_204: Not I; Site_205: Not I; Site_206: Not I; Site_207:
Not I; Site_208: Not I; Site_209: Not I; Site_210: Not I; Site_211:
Not I; Site_212: Not I; Site_213: Not I; Site_214: Not I; Site_215:
Not I; Site_216: Not I; Site_217: Not I; Site_218: Not I; Site_219:
Not I; Site_220: Not I; Site_221: Not I; Site_222: Not I; Site_223:
Not I; Site_224: Not I; Site_225: Not I; Site_226: Not I; Site_227:
Not I; Site_228: Not I; Site_229: Not I; Site_230: Not I; Site_231:
Not I; Site_232: Not I; Site_233: Not I; Site_234: Not I; Site_235:
Not I; Site_236: Not I; Site_237: Not I; Site_238: Not I; Site_239:
Not I; Site_240: Not I; Site_241: Not I; Site_242: Not I; Site_243:
Not I; Site_244: Not I; Site_245: Not I; Site_246: Not I; Site_247:
Not I; Site_248: Not I; Site_249: Not I; Site_250: Not I; Site_251:
Not I; Site_252: Not I; Site_253: Not I; Site_254: Not I; Site_255:
Not I; Site_256: Not I; Site_257: Not I; Site_258: Not I; Site_259:
Not I; Site_260: Not I; Site_261: Not I; Site_262: Not I; Site_263:
Not I; Site_264: Not I; Site_265: Not I; Site_266: Not I; Site_267:
Not I; Site_268: Not I; Site_269: Not I; Site_270: Not I; Site_271:
Not I; Site_272: Not I; Site_273: Not I; Site_274: Not I; Site_275:
Not I; Site_276: Not I; Site_277: Not I; Site_278: Not I; Site_279:
Not I; Site_280: Not I; Site_281: Not I; Site_282: Not I; Site_283:
Not I; Site_284: Not I; Site_285: Not I; Site_286: Not I; Site_287:
Not I; Site_288: Not I; Site_289: Not I; Site_290: Not I; Site_291:
Not I; Site_292: Not I; Site_293: Not I; Site_294: Not I; Site_295:
Not I; Site_296: Not I; Site_297: Not I; Site_298: Not I; Site_299:
Not I; Site_300: Not I; Site_301: Not I; Site_302: Not I; Site_303:
Not I; Site_304: Not I; Site_305: Not I; Site_306: Not I; Site_307:
Not I; Site_308: Not I; Site_309: Not I; Site_310: Not I; Site_311:
Not I; Site_312: Not I; Site_313: Not I; Site_314: Not I; Site_315:
Not I; Site_316: Not I; Site_317: Not I; Site_318: Not I; Site_319:
Not I; Site_320: Not I; Site_321: Not I; Site_322: Not I; Site_323:
Not I; Site_324: Not I; Site_325: Not I; Site_326: Not I; Site_327:
Not I; Site_328: Not I; Site_329: Not I; Site_330: Not I; Site_331:
Not I; Site_332: Not I; Site_333: Not I; Site_334: Not I; Site_335:
Not I; Site_336: Not I; Site_337: Not I; Site_338: Not I; Site_339:
Not I; Site_340: Not I; Site_341: Not I; Site_342: Not I; Site_343:
Not I; Site_344: Not I; Site_345: Not I; Site_346: Not I; Site_347:
Not I; Site_348: Not I; Site_349: Not I; Site_350: Not I; Site_351:
Not I; Site_352: Not I; Site_353: Not I; Site_354: Not I; Site_355:
Not I; Site_356: Not I; Site_357: Not I; Site_358: Not I; Site_359:
Not I; Site_360: Not I; Site_361: Not I; Site_362: Not I; Site_363:
Not I; Site_364: Not I; Site_3
```

Best Local Similarity 99.8%; Pred. NO. 0.00e+00;
Matches 459; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

Db 1 AAAGTTTATACAGATATAAAGTTTATCTGAAATATTTACAGTGTGGTTAAAGCAA 60
|||||
Cp 1986 AAAGTTTATACAGATATAAAGTTTATCTGAAATATTTACAGTGTGGTTAAAGCAA 1927
|||||
Db 61 TATTTTCAACACTTTAAAGCTAAACTACTATGTATATATACAGTACAGCTACAAATGGGTT 120
|||||
Cp 1926 TATTTTCAACACTTTAAAGCTAAACTACTATGTATATATACAGTACAGCTACAAATGGGTT 1867
|||||
Db 121 TAAATTGCAAAAGTTAACTAGAAATGTTTTAAACAGGCTTAAAGTACTCAAGTCAATT 180
|||||
Cp 1866 TAAATTGCAAAAGTTAACTAGAAATGTTTTAAACAGGCTTAAAGTACTCAAGTCAATT 1807
|||||
Db 181 ATAAATTTATATCTTTTGCCTTTTACTTGAAGAAATCATGCTATAGAAATGGTTAATGT 240
|||||
Cp 1806 ATAAATTTATATCTTTTGCCTTTTACTTGAAGAAATCATGCTATAGAAATGGTTAATGT 1747
|||||
Db 241 GCTTCTAATAATGGAAGTATTTAGTGTGAATGTGATACATGTAAACAGTTTAAAGTTCCTC 300
|||||
Cp 1746 GCTTCTAATAATGGAAGTATTTAGTGTGAATGTGATACATGTAAACAGTTTAAAGTTCCTC 1687
|||||
Db 301 ATTGAAGTATAAATGATGAATTCGTTGTAAGACTTAGACACTGAGTCTCAGTCTCGAG 360
|||||
Cp 1686 ATTGAAGTATAAATGATGAATTCGTTGTAAGACTTAGACACTGAGTCTCAGTCTCGAG 1628
|||||
Db 361 CTGATGAAGATGTTGAGATACACGCCAGCTTATCTCAACAGGTTTGTGACCCACAAAGT 420
|||||
Cp 1627 CTGATGAAGATGTTGAGATACACGCCAGCTTATCTCAACAGGTTTGTGACCCACAAAGT 1568
|||||
Db 421 TTGGCCACACAGAGAAATTTGAAGCAATTTGCATGTTATGA 460
|||||
Cp 1567 TTGGCCACACAGAGAAATTTGAAGCAATTTGCATGTTATGA 1528
|||||

RESULT 8
LOCUS AA449300 443 bp mRNA EST 04-JUN-1997
DEFINITION zx08c06.s1 Soares total fetus Nb2HF8 9w Homo sapiens cDNA clone 785866 3', mRNA sequence.
ACCESSION AA449300
NID G2163149
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 443)
AUTHORS Hallier,L., Allen,M., Bowles,L., Dubuque,T., Geisel,G., Jost,S., Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M., Martin,J., Moore,B., Schellenberg,K., Steptoe,M., Tan,F., Theising,B., White,Y., Wylie,T., Waterston,R. and Wilson,R.
TITLE Washu-Merck EST Project 1997
JOURNAL Unpublished (1997)
COMMENT Contact: Wilson RK
Washu-Merck EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available royalty-free through LLNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
Seq primer: -41m13 fwd. ET from Amersham
High quality sequence stop: 412.
Location/Qualifiers
1. .443
/organism="Homo sapiens"
/note="Vector: pT73D-Pac (Pharmacia) with a modified polylinker; Site_1: Not 1; Site_2: Eco RI; 1st strand cDNA was prepared from mRNA obtained from pooled 8-9 week

(total) fetus material with a Not I - oligo(GT) primer [5', TGTTCACCAATCTGAAGTGGAGCGCCGCTTAATTTTCTTTTCTTTT 3']. Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT73 vector. Library went through one round of normalization, and was constructed by Bento Soares and M. Fatima Bonaudo." /db_xref="taxon:9606"
/clone="785866"
/clone_lib="Soares total fetus Nb2HF8 9w"
/dev_stage="8-9 weeks"
/lab_host="DH10B"
complement(<1..>443)
/db_xref="GDB:5983485"

BASE COUNT 153 a 47 c 83 g 160 t
ORIGIN

Query Match 21.4%; Score 434; DB 7; Length 443;
Best Local Similarity 99.5%; Pred. NO. 0.00e+00;
Matches 441; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

Db 1 TTTTCTTATTTTATTTAGGTAACACAGGATGTAAGGTTTATATACAGCAATATAATGTTTA 60
|||||
Cp 2017 TTTTCTTATTTTATTTAGGTAACACAGGATGTAAGGTTTATATACAGCAATATAATGTTTA 1958
|||||
Db 61 TCTGAAATATTTACAGTGTGGTTAAAGCAATATTTTACAACTTTTAAAGGTAAACTAC 120
|||||
Cp 1957 TCTGAAATATTTACAGTGTGGTTAAAGCAATATTTTACAACTTTTAAAGGTAAACTAC 1898
|||||
Db 121 TATGTATATTACAGGTAAGTCAATGGGTTTAAATTTGCAAAAGTTTAAGTAAGAAATGTT 180
|||||
Cp 1897 TATGTATATTACAGGTAAGTCAATGGGTTTAAATTTGCAAAAGTTTAAGTAAGAAATGTT 1838
|||||
Db 181 TTAACAAGGCTTAAAGTACTCAAGTCAATTAATAAATTTATATCTTTGCGCTTTTACTT 240
|||||
Cp 1837 TTAACAAGGCTTAAAGTACTCAAGTCAATTAATAAATTTATATCTTTGCGCTTTTACTT 1778
|||||
Db 241 GAAGAAATCATGCTATAGAAATGGTTAAATGTGCTTCTAATAAATGGAAGTATTGTAGCTG 300
|||||
Cp 1777 GAAGAAATCATGCTATAGAAATGGTTAAATGTGCTTCTAATAAATGGAAGTATTGTAGCTG 1718
|||||
Db 301 GAATGTGATACATGTAAACAGTTTAAGTTCCTTCCCATTAAGGTATAAATGATGAATGTTGG 360
|||||
Cp 1717 GAATGTGATACATGTAAACAGTTTAAGTTCCTTCCCATTAAGGTATAAATGATGAATGTTGG- 1659
|||||
Db 361 TAAGACTTAGACACTGGGTCCTCAGTCTGGAGCTGTGAAGATGTTGAGATACAGCCAGC 420
|||||
Cp 1658 TAAGACTTAGACACTGGGTCCTCAGTCTGGAGCTGTGAAGATGTTGAGATACAGCCAGC 1599
|||||
Db 421 TTTATCTCAACAGGGTTTGTGAC 443
|||||
Cp 1598 TTTATCTCAACAGGGTTTGTGAC 1576
|||||

RESULT 9
LOCUS AA922948 482 bp mRNA EST 09-JUN-1998
DEFINITION ch10h11.s1 NCI CGAP Co8 Homo sapiens cDNA clone IMAGE:1457445 3', similar to TR:O08862 O08862 SECRETED FRIZZLED RELATED PROTEIN SFRP-2. [2] TR:P97299 ;, mRNA sequence.
ACCESSION AA922948
NID G3070257
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 482)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.

Tel: (301) 496-1550
 Email: Robert.Strausberg@nih.gov
 Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D.
 CDNA Library Preparation: M. Bento Soares, Ph.D.
 CDNA Library Arrayed by: Greg Lennon, Ph.D.
 DNA Sequencing by: Washington University Genome Sequencing Center
 Clone distribution: NCI-CCAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
 www-bio.llnl.gov/bbrp/image/image.html

Insert Length: 870 Std Error: 0.00
 Seq primer: -40ml3 fwd. ET from Amersham
 High quality sequence stop: 377.

FEATURES

1. 482
 Location/Qualifiers
 /organism="Homo sapiens"
 /note="Organ: colon; Vector: p773D-Pac (Pharmacia) with a modified polylinker; 1st strand cDNA was prepared from colon adenocarcinoma, and was then primed with a Not I - oligo(dT) primer. Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified p773 vector. Library is normalized. Library was constructed by Bento Soares and M. Fatima Bonaldo."
 /db_xref="taxon:9606"
 /clone="IMAGE:1457445"
 /clone_lib="NCI-CCAP_C08"
 /tissue_type="adenocarcinoma"
 /lab_host="DH10B"

BASE COUNT 103 a 119 c 127 g 133 t

Query Match 21.4%; Score 433; DB 18; Length 482;
 Best Local Similarity 99.8%; Pred. No. 0.00e+00;
 Matches 434; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Db 10 AAAATATTTCATGATTTTATAGTTTGAATATTTCTACAGATTCGGGTGGGCTTTTCCTT 69
 Cp 1406 AAAATATTTCATGATTTTATAGTTTGAATATTTCTACAGATTCGGGTGGGCTTTTCCTT 1347

Db 70 TAGGTGAAACAGCTATCCACTCCTGTGCGCTTATAAATCAGAGAAATGCTGGGATGCAA 129
 Cp 1346 TAGGTGAAACAGCTATCCACTCCTGTGCGCTTATAAATCAGAGAAATGCTGGGATGCAA 1287

Db 130 AGTGCAGGAGCGGGGAGGAGTCCGAGCTGAGACTGGAGCAGCTAGGAGTGTGCTT 189
 Cp 1286 AGTGCAGGAGCGGGGAGGAGTCCGAGCTGAGACTGGAGCAGCTAGGAGTGTGCTT 1227

Db 190 GGGAGACGGAGCTGAGATCCCGAGCAGAAATGGTCAGCGCTCTGGAGCAGGCGCTG 249
 Cp 1226 GGGAGACGGAGCTGAGATCCCGAGCAGAAATGGTCAGCGCTCTGGAGCAGGCGCTG 1167

Db 250 TCGGAGCCATCAGGATGCGGGAGCTAGCAGCTGCGAATGCTGCGGAGATGCGC 309
 Cp 1166 TCGGAGCCATCAGGATGCGGGAGCTAGCAGCTGCGAATGCTGCGGAGATGCGC 1107

Db 310 TTGAATCTCTGCGCCCTTCTGCCACCGCTTCACCGAGGTGATCACCAGCTCCGCCACCC 369
 Cp 1106 TTGAATCTCTGCGCCCTTCTGCCACCGCTTCACCGAGGTGATCACCAGCTCCGCCACCC 1047

Db 370 TGTTTCGTCCATGACAGATAGGCGGCTTGATGCTGCTTCATCTCCTCAGAGGTGCAC 429
 Cp 1046 TGTTTCGTCCATGACAGATAGGCGGCTTGATGCTGCTTCATCTCCTCAGAGGTGCAC 987

Db 430 TCGAAGCTGCTTTG 444
 Cp 986 TCGAAGCTGCTTTG 972

RESULT 10
 LOCUS AA431859 427 bp mRNA EST 22-MAY-1997
 DEFINITION zw51g08.s1 Soares total fetus Nb2Hf8 9w Homo sapiens cDNA clone

ACCESSION AA431859
 NID G2115567
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryotes; mitochondrial eukaryotes; Metazoa; Chordata; Vertebrata; Mammalia; Euthera; Primates; Catarrhini; Hominiidae; Homo.

REFERENCE

AUTHORS 1 (bases 1 to 427)
 Hillier, L., Allen, M., Bowles, L., Dubuque, T., Geisel, G., Jost, S., Kucaba, T., Lacy, M., Le, N., Lennon, G., Marra, M., Martin, J., Moore, B., Schellenberg, K., Steptoe, M., Tan, F., Theising, B., White, Y., Wylie, T., Waterston, R. and Wilson, R.
 WashU-Merck EST Project 1997
 Unpublished (1997)

TITLE

JOURNAL
 COMMENT

Contact: Wilson RK
 WashU-Merck EST Project
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: est@watson.wustl.edu
 This clone is available royalty-free through LLNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
 Seq primer: -40ml3 fwd. ET from Amersham
 High quality sequence stop: 214.
 Location/Qualifiers
 1. 427

FEATURES

source
 /organism="Homo sapiens"
 /note="Vector: p773D-Pac (Pharmacia) with a modified polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA was prepared from mRNA obtained from pooled 8-9 week (total) fetus material with a Not I oligo(dT) primer [5' TGTTACCAATCTGAAGTGGGAGCGGCTTAATTTTTTTTTTTT 3']
 Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified p773 vector. Library went through one round of normalization, and was constructed by Bento Soares and M. Fatima Bonaldo."
 /db_xref="taxon:9606"
 /clone="773630"
 /dev_stage="8-9 weeks"
 /lab_host="DH10B"
 complement(<1..>427) 148 t

mRNA
 BASE COUNT 154 a 46 c 79 g 148 t
 ORIGIN

Query Match 21.0%; Score 425; DB 5; Length 427;
 Best Local Similarity 99.8%; Pred. No. 0.00e+00;
 Matches 426; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Db 1 TTTTATAGTAAACAGGATGTAAGTTTATATACAGAATAATATGTTTCTCAATA 60
 Cp 2008 TTTTATAGTAAACAGGATGTAAGTTTATATACAGAATAATATGTTTCTCAATA 1949

Db 61 TTTACAGTGTGGTTAAAGCAATATTTTACAACCTTTTAAAGGTAACCTATATATAT 120
 Cp 1948 TTTACAGTGTGGTTAAAGCAATATTTTACAACCTTTTAAAGGTAACCTATATATAT 1889

Db 121 TACAGGTAAAGTACAAATGGGTTTAAATTTGCAAAAGTAAAGTAAGAAATGTTTAAACAAG 180
 Cp 1888 TACAGGTAAAGTACAAATGGGTTTAAATTTGCAAAAGTAAAGTAAGAAATGTTTAAACAAG 1829

Db 181 GCTTAAAGTACTCAAGTCAATTAATAATTTATATATATATATATATATATATATATAT 240
 Cp 1828 GCTTAAAGTACTCAAGTCAATTAATAATTTATATATATATATATATATATATATATAT 1769

Db 241 ATGCTATAGAAATGTTTAAATGCTTCTTAATAAATGGAAGTATTTAGCTGGAATGTGAT 300
 Cp 1768 ATGCTATAGAAATGTTTAAATGCTTCTTAATAAATGGAAGTATTTAGCTGGAATGTGAT 1709

```

Db 301 ACATGTAACAGTTTAAAGTCCCATGAAAGGTATATAAAATGATGAATTTGTTAAGACTTAG 360
|||||
Cp 1708 ACATGTAACAGTTTAAAGTCCCATGAAAGGTATATAAAATGATGAATTTGTTAAGACTTAG 1649
|||||
Db 361 ACACGTAGTCTCAGTCTGGAGCTGATGAAGATGTTGAGATACAGCCAGCTTTATCTCAA 420
|||||
Cp 1648 ACACGTAGTCTCAGTCTGGAGCTGATGAAGATGTTGAGATACAGCCAGCTTTATCTCAA 1589
|||||
Db 421 CAGGGTT 427
|||||
Cp 1588 CAGGGTT 1582
|||||

RESULT 11
LOCUS AA927991 518 bp mRNA EST 22-APR-1998
DEFINITION on47d08.s1 NCI_CGAP_C08 Homo sapiens cDNA clone IMAGE:155923 3'
similar to TR:008862 008862 SECRETED FRIZZLED RELATED PROTEIN
SRFP-2. [2] TR:P97299 ;, mRNA sequence.
ACCESSION AA927991
NID g3076735
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 518)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert.Strausberg@nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: M. Bento Soares, Ph.D.
DNA Library Arrayed by: Greg Lennon, Ph.D.
Sequence generated by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html

Seg primer: -40m13 fwd. ET from Amersham
High quality sequence stop: 194.
FEATURES
source
Location/Qualifiers
1..518
/organism="Homo sapiens"
/note="Organ: colon; Vector: pT7T3D-Pac (Pharmacia) with a
modified polylinker; 1st strand cDNA was prepared from
colon adenocarcinoma, and was then primed with a Not I -
oligo(dT) primer. Double-stranded cDNA was ligated to Eco
RI adaptors (Pharmacia), digested with Not I and cloned
into the Not I and Eco RI sites of the modified pT7T3
vector. Library is normalized. Library was constructed by
Bento Soares and M. Fatima Bonaldo."
/db_xref="taxon:9606"
/clone="IMAGE:1559823"
/clone_lib="NCI_CGAP_C08"
/tissue_type="adenocarcinoma"
/lab_host="DH10B"
BASE COUNT 112 a 123 c 141 g 142 t
ORIGIN
Query Match 20.7%; Score 420; DB 17; Length 518;
Best Local Similarity 98.8%; Pred.No. 0.00e+00;
Matches 425; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Db 10 AAAATATTTCATGATTTTATTAGTTTGAATATTTCTACAAGATTCGGGTGGGCTTTTCCTT 69
|||||
Cp 1406 AAAATATTTCATGATTTTATTAGTTTGAATATTTCTACAAGATTCGGGTGGGCTTTTCCTT 1347
|||||

```

```

Db 70 TAGTGAAGAACAGCTATCCACTCTCTGTGGCTTTATAACTCAGGAATGCTGGGGATGCAA 129
|||||
Cp 1346 TAGTGAAGAACAGCTATCCACTCTCTGTGGCTTTATAACTCAGGAATGCTGGGGATGCAA 1287
|||||
Db 130 AGTGTCAAAAGGAGGGGGAAGCTGCCAGGCTGAGACTGGAGCAGCTAGGAGTGTGCTT 189
|||||
Cp 1286 ACCTGTCAAAAGGAGGGGGAAGCTGCCAGGCTGAGACTGGAGCAGCTAGGAGTGTGCTT 1227
|||||
Db 190 GGGGACGGGAGCTGAGATCCCGGAGCAGAAATGGTCAGCCGTGCTGTGAGCAGGCCTG 249
|||||
Cp 1226 GGGGACGGGAGCTGAGATCCCGGAGCAGAAATGGTCAGCCGTGCTGTGAGCAGGCCTG 1167
|||||
Db 250 TCGGAGCCATCAGGATGCGGGGACTAGCAGCTTTCGGGATGCTGCGGGAGATGCGC 309
|||||
Cp 1166 TCGGAGCCATCAGGATGCGGGGACTAGCAGCTTTCGGGATGCTGCGGGAGATGCGC 1107
|||||
Db 310 TTGAACCTCTCTGCCCCCTTCTGCCACCGCTTCGACGAGGTATCACCAGCTCCCAACC 369
|||||
Cp 1106 TTGAACCTCTCTGCCCCCTTCTGCCACCGCTTCGACGAGGTATCACCAGCTCCCAACC 1047
|||||
Db 370 TGTTCCTGCTCCCATGACCATAGGCGCGCTTATGTGTTTCATCTACTACACAGGTGCAC 429
|||||
Cp 1046 TGTTCCTGCTCCCATGACCATAGGCGCGCTTATGTGTTTCATCTACTACACAGGTGCAC 987
|||||
Db 430 TGAAGCTGT 439
|||||
Cp 986 TGCAAGCTGT 977
|||||

RESULT 12
LOCUS N64004 446 bp mRNA EST 01-MAR-1996
DEFINITION YY64d09.s1 Homo sapiens cDNA clone 278321 3'.
ACCESSION N64004
NID g1211833
KEYWORDS EST.
SOURCE human clone=278321 primer=ml3 -40 forward library=Soares multiple
sclerosis 2NBHSP vector=pT7T3D (Pharmacia) with a modified
polylinker V-TYPE: phagemid host=DH10B (ampicillin resistant)
Rsitel=Not I Rsitel2=Eco RI 46 year old male. 1st strand cDNA was
primed with a Not I - oligo(dT) primer
[5'-TGTTACCAATCTGAGTGGGAGGCGGCGCATTTTTCCTTTTTCCTTTT-3'],
double-stranded cDNA was size selected, ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I and Eco
RI sites of a modified pT7T vector (Pharmacia). Library went
through one round of normalization to a Cot = 5. Library
constructed by Bento Soares and M.Fatima Bonaldo. RNA from 4
multiple sclerosis lesions from one patient was kindly provided by
Dr. Kevin G. Becker (NINDS/NIH).
ORGANISM Homo sapiens
Eukaryota; Metazoa; Eumetazoa; Bilateria; Coelomata;
Deuterostomia; Chordata; Vertebrata; Gnathostomata; Osteichthyes;
Sarcopterygii; Chonata; Tetrapoda; Amniota; Mammalia; Theria;
Eutheria; Archonta; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 446)
REFERENCE Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,
Holman,M., Huitman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,
Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F.,
Trevaskis,E., Waterston,R., Williamson,A., Wohldmann,P. and
Wilson,R.
The WashU-Merck EST Project
Unpublished (1995)
Contact: Wilson RK
WashU-Merck EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@wustl.wustl.edu
High quality sequence stops: 225
Source: IMAGE Consortium, LLNL
This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.

```


SOURCE human.
ORGANISM Homo sapiens
Eukaryotes; mitochondrial eukaryotes; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.

REFERENCE 1 (bases 1 to 409)
AUTHORS Hillier, L., Allen, M., Bowles, L., Dubuque, T., Geisel, G., Jost, S., Kucaba, T., Lacy, M., Le, N., Lennon, G., Marra, M., Martin, J., Moore, B., Schellenberg, K., Steptoe, M., Tan, F., Theising, B., White, Y., Wylie, T., Waterston, R., and Wilson, R.
TITLE WashU-Merck EST Project 1997
JOURNAL Unpublished (1997)
COMMENT Contact: Wilton RK
WashU-Merck EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available royalty-free through LLNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
Seq primer: -4lnl3 fwd. Et from Amersham
High quality sequence stop: 408.

FEATURES
source
1..409
/organism="Homo sapiens"
/note="vector: p7T3D-Pac (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was prepared from mRNA obtained from pooled 8-9 week (total) fetus material with a Not I - oligo(dT) primer [5' TGTTACCAATCTGAAGTGGAGCGCGCTTAATTTTTTTTTTTT 3']
Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT73 vector. Library went through one round of normalization, and was constructed by Bento Soares and M. Fatima Bonaldo."
/db_xref="taxon:9606"
/clone_lib="Soares total fetus NB2HF8 9w"
/dev_stage="8-9 weeks"
/lab_host="DH10B"
complement(<1..>409)
/db_xref="GDB:5963421" 144 t

mRNA
148 a 42 c 75 g 144 t

BASE COUNT
ORIGIN

Query Match 20.2%; Score 409; DB 7; Length 409;
Best Local Similarity 100.0%; Pred. No. 0.00e+00;
Matches 409; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 TTTTCTTTAGTAAACAGGATGTAAGTTTATATACAGAATATAATGTTTATCTGAAA 60
Cp 2010 TTTTCTTTAGTAAACAGGATGTAAGTTTATATACAGAATATAATGTTTATCTGAAA 1951
Db 61 TATTACAGTGTGGTTAAAGCATATTTTACAACTTTTAAAGGTAACACTACTATGAT 120
Cp 1950 TATTACAGTGTGGTTAAAGCATATTTTACAACTTTTAAAGGTAACACTACTATGAT 1891
Db 121 ATTACAGTAACTACAACTGGTTTAAATTTGCAAAAGTTAAAGTAAGAAATGTTTAAACA 180
Cp 1890 ATTACAGTAACTACAACTGGTTTAAATTTGCAAAAGTTAAAGTAAGAAATGTTTAAACA 1831
Db 181 AGGCTTAAAGTAACTACAACTGGTTTAAATTTGCAAAAGTTAAAGTAAGAAATGTTTAAACA 240
Cp 1830 AGGCTTAAAGTAACTACAACTGGTTTAAATTTGCAAAAGTTAAAGTAAGAAATGTTTAAACA 1771
Db 241 TCATGCTTACAAATGTTTAAATGCTTCTTAATAAATGGAAGTATGCTAGCTGAATGTC 300
Cp 1770 TCATGCTTACAAATGTTTAAATGCTTCTTAATAAATGGAAGTATGCTAGCTGAATGTC 1711
Db 301 ATACATGTAACAGTTAAAGTTCCTCCATTTGAGGTTAAAGTATAAATGATGAATGTTGTAAGACTT 360

Cp 1710 ATACATGTAACAGTTTAAAGTTCCTCCATTTGAAGGTATATAAATGATGAATGTTGTAAGACTT 1651
Db 361 AGACACTGAGTCTCAGTCTCGAGCTGATGAAGATGTTGAGATAACAGCC 409
Cp 1650 AGACACTGAGTCTCAGTCTCGAGCTGATGAAGATGTTGAGATAACAGCC 1602

RESULT 15
LOCUS N32415 445 bp mRNA EST 10-JAN-1996
DEFINITION YW83C06.s1 Homo sapiens cDNA clone 258826 3'
ACCESSION N32415
NID g1152814
KEYWORDS EST.
SOURCE human clone=258826 primer=ml3 -40 forward library=Soares placenta 8to9weeks 2NbHP8to9W vector=pT7T3D (Pharmacia) with a modified polylinker host=DH10B (ampicillin resistant) Rsitel=Not I Rsitel2=Eco RI two placentae: one from 8 weeks and another from 9 weeks post conception. 1st strand cDNA was primed with a Not I - oligo(dT) primer
[5'-TGTTACCAATCTGAAGTGGAGCGCGCATTTTTTTTTTTTTTTT-3'], double-stranded cDNA was size selected, ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of a modified pT73 vector (Pharmacia). Library constructed by Bento Soares and M. Fatima Bonaldo.

ORGANISM
Homo sapiens
Eukaryotes; Metazoa; Eumetazoa; Bilateria; Coelomata; Deuterostomia; Chordata; Vertebrata; Gnathostomata; Osteichthyes; Sarcopterygii; Choanata; Tetrapoda; Amniota; Mammalia; Theria; Eutheria; Archonta; Primates; Catarrhini; Hominoidea; Homo. 1 (bases 1 to 445)
Hillier, L., Clark, N., Dubuque, T., Elliston, K., Hawkins, M., Holman, M., Hultman, M., Kucaba, T., Le, M., Lennon, G., Marra, M., Parsons, J., Rifkin, L., Rohlfing, T., Soares, M., Tan, F., Trevaskis, E., Waterston, R., Williamson, A., Wohlmann, P., and Wilson, R.
TITLE The WashU-Merck EST Project
JOURNAL Unpublished (1995)
COMMENT Contact: Wilton RK
WashU-Merck EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
High quality sequence stops: 314
Source: IMAGE Consortium, LLNL
This clone is available royalty-free through LLNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.

FEATURES
source
1..445
/organism="Homo sapiens"
/clone="258826"
<1..>445

BASE COUNT
ORIGIN
156 a 55 c 87 g 147 t

Query Match 20.1%; Score 408; DB 20; Length 445;
Best Local Similarity 98.6%; Pred. No. 0.00e+00;
Matches 434; Conservative 0; Mismatches 2; Indels 4; Gaps 4;

Db 1 AAACAGAGTGAAGTCTTATATACAGAATATAATGTTTATCTGAAATATTACAGTGT 60
Cp 1998 AAACAGAGTGAAGTCTTATATACAGAATATAATGTTTATCTGAAATATTACAGTGT 1939
Db 61 TGGTTAAACCAATATTTTACAACTTTTAAAGGTAACACTACTATGTTTATACAGTAA 120
Cp 1938 TGGTTAAACCAATATTTTACAACTTTTAAAGGTAACACTACTATGTTTATACAGTAA 1879
Db 121 CTACATGGGTTTAAATTTGCAAAAGTTAAAGTAAGAAATGTTTAAACAGGCTTAAAGTA 180
Cp 1878 CTACATGGGTTTAAATTTGCAAAAGTTAAAGTAAGAAATGTTTAAACAGGCTTAAAGTA 1819

Db 181 CTCAGTCAATTATAAAATTATATCTTTTGCCTTTTACTTGAAGAAATCATGCTATAGA 240
 Cp 1818 CTCAGTCAATTATAAAATTATATCTTTTGCCTTTTACTTGAAGAAATCATGCTATAGA 1759
 Db 241 AATGGTTAATGCTCTCTAATAATGAAGTATTGTAGCTGGGAATGTGATACATGTAACA 300
 Cp 1758 AATGGTTAATGCTCTCTAATAATGAAGTATTGTAGCTGGGAATGTGATACATGTAACA 1699
 Db 301 GTTTAAGTTCCCATTTGAAGGTATAAAATGATGAATTGTTGTAAGACTTAGACACTGGAGT 360
 Cp 1698 GTTTAAGTTCCCATTTGAAGGTATAAAATGATGAATTGTTGTAAGACTTAGACACTG-AGT 1640
 Db 361 CTCAGTCTGGGAGCTGATGAAGATGTTGAGATAACAGCCAGCTTTATCCCCAACAGGGGT 420
 Cp 1639 CTCAGTCTGG-AGCTGATGAAGATGTTGAGATAACAGCCAGCTTTATCTC-AACAGGG-T 1583
 Db 421 TTGTGACCCACAAGTTTGG 440
 Cp 1582 TTGTGACCCACAAGTTTGG 1563

Search completed: Wed Aug 5 07:34:29 1998
 Job time : 4465 secs.

THIS PAGE BLANK (section)

Result No.	Query	Score	Match	Length	DB	ID	Description	Pred. No.
1	396	17.7	572	27	W31273	Mouse frizzled-7 prot	2.72e-28	
2	394	17.6	685	27	W31274	Mouse frizzled-8 prot	4.35e-28	
3	374	16.7	585	27	W31271	Human frizzled-5 prot	4.69e-26	
4	369	16.5	537	27	W31270	Mouse frizzled-4 prot	1.51e-25	
5	357	16.0	525	27	W31269	Caenorhabditis frizsl	2.45e-24	
6	355	15.9	694	27	W31267	Drosophila frizzled-2	3.92e-24	
7	317	14.2	666	27	W31268	Mouse frizzled-3 prot	2.56e-20	
8	284	12.7	709	27	W31272	Mouse frizzled-6 prot	4.86e-17	
9	148	6.6	581	14	R74187	Chick p75	3.74e-04	
10	121	5.6	605	14	R74186	Chick p78	4.55e-02	
11	124	5.4	604	28	W35946	Human netrin-1.	8.16e-02	
12	105	4.7	529	14	R74188	Mouse p78.	1.70e+00	
13	100	4.5	732	25	W26642	Human RECK cancer-inh	4.26e+00	
14	99	4.4	804	29	F35368	H. pylori ORF 09ap114	5.11e+00	
15	96	4.3	868	25	W26610	Rat muscle-specific k	8.79e+00	
16	96	4.3	868	25	W26507	Rat Dmk receptor.	8.79e+00	
17	95	4.3	869	25	W26611	Human musc-specific	1.05e+01	
18	95	4.3	869	25	W26506	Human Dmk receptor.	1.05e+01	

QY 44 PANLQCHGIEYQNMRLPNLLGHETMKVEYLEQAGAWIPLVMKQCHDPTKKFLCSLFAFVC 103
Db 111 -tylqaiopqslqkqararqgqcalmknkfgwperllcenfsvhgageiaygqntsd 167
QY 104 LDDLETIQCHSLCVQVQKDRCAPVNSAFGFPWMDLECDREP-QDN-DLCIPLASSD 159

RESULT 2
ID W31274 standard; Protein; 585 AA.
AC W31274;
DE 27-APR-1998 (first entry)
DT Mouse frizzled-8 protein Mfz8 (Wnt receptor).
KW Wnt receptor; mouse frizzled-8 protein; Mfz8 gene;
KW signal transduction; cancer; cell growth; cell proliferation.
OS Mus musculus.
PN W09739357-A1.
PD 23-OCT-1997.
PF 11-APR-1997; U06049.
PR 12-APR-1996; US-015307.
PA (UYJO) UNIV JOHNS HOPKINS.
PA (STRD) UNIV LELAND STANFORD JUNIOR.
PI Andrew D, Bhanot P, Brink M, Hsieh J, Nathans J,
PI Nusse R, Samos CH, Wangy;
DR N-PSDB; T89889.
PT Identification of Wnt receptor binding modulators - useful for
PT treatment of cancer and growth, development or proliferation related
PT disorders
PS Disclosure; Page 48-50; 61pp; English.
CC This protein comprises the mouse transmembrane receptor,
CC frizzled-8 (Mfz8), encoded by the Mfz8 gene (see T89892). It is
CC an example of a Wnt receptor. Other novel frizzled family members
CC have been identified in human, mouse and Caenorhabditis elegans
CC (see W31268-74) and are considered also to be Wnt receptors. Wnt
CC receptors can be used in a novel, claimed method of screening for
CC compounds which modulate the binding of a Wnt polypeptide (secreted
CC proteins involved in cell-to-cell signalling) to a Wnt receptor.
CC Wnt is involved in (mammary) cancer and other processes involving
CC growth, development and proliferation (both normal and abnormal).
CC Modulators identified by the claimed method are useful for
CC treatment of diseases related to these conditions.
SQ Sequence 585 AA;

Query Match 17.6%; Score 394; DB 27; Length 685;
Best Local Similarity 43.8%; Pred. NO. 4.35e-28;
Matches 56; Conservative 23; Mismatches 43; Indels 6; Gaps 6;

Db 40 vplckgignytmnqfnhdtqdeaglevhgf-piveiqcsdtkfflcsmytpicle 98
QY 47 LQCHGIEYQNMRLPNLLGHETMKVEY-LEQAGAWIPLVMKQCHDPTKKFLCSLFAFVC 105

Db 99 dykxlpccrsvcrakagcaplmrvgfawpdmrcdripegqnpdltcmndyartd-lt 157
QY 106 DLDETIQCHSLCVQVQKDRCAPVNSAFGFPWMDLECDREP-QDN-D-LCIPLASSDHL 162

Db 158 taapsppr 165
QY 163 PATEAPK 170

RESULT 3
ID W31271 standard; Protein; 585 AA.
AC W31271;
DE 27-APR-1998 (first entry)
DT Human frizzled-5 protein Mfz5 (Wnt receptor).
KW Wnt receptor; human frizzled-5 protein; Mfz5 gene;
KW signal transduction; cancer; cell growth; cell proliferation.
OS Homo sapiens.
PN W09739357-A1.
PD 23-OCT-1997.
PF 11-APR-1997; U06049.
PR 12-APR-1996; US-015307.
PA (UYJO) UNIV JOHNS HOPKINS.

QY 44 PANLQCHGIEYQNMRLPNLLGHETMKVEYLEQAGAWIPLVMKQCHDPTKKFLCSLFAFVC 103
PI Andrew D, Bhanot P, Brink M, Hsieh J, Nathans J,
PI Nusse R, Samos CH, Wangy;
DR N-PSDB; T89889.
PT Identification of Wnt receptor binding modulators - useful for
PT treatment of cancer and growth, development or proliferation related
PT disorders
PS Disclosure; Page 34-35; 61pp; English.
CC This protein comprises the mouse transmembrane receptor,
CC frizzled-4 (Mfz4), encoded by the Mfz4 gene (see T89888). It is
CC an example of a Wnt receptor. Other novel frizzled family members
CC have been identified in human, mouse and Caenorhabditis elegans
CC (see W31268-74) and are considered also to be Wnt receptors. Wnt
CC receptors can be used in a novel, claimed method of screening for
CC compounds which modulate the binding of a Wnt polypeptide (secreted
CC proteins involved in cell-to-cell signalling) to a Wnt receptor.
CC Wnt is involved in (mammary) cancer and other processes involving
CC growth, development and proliferation (both normal and abnormal).
CC Modulators identified by the claimed method are useful for

PA (STRD) UNIV LELAND STANFORD JUNIOR.
PI Andrew D, Bhanot P, Brink M, Hsieh J, Nathans J,
PI Nusse R, Samos CH, Wangy;
DR N-PSDB; T89889.
PT Identification of Wnt receptor binding modulators - useful for
PT treatment of cancer and growth, development or proliferation related
PT disorders
PS Disclosure; Page 37-39; 61pp; English.
CC This protein comprises the human transmembrane receptor,
CC frizzled-5 (Hfz5), encoded by the Hfz5 gene (see T89889). It is
CC an example of a Wnt receptor. Other novel frizzled family members
CC have been identified in human, mouse and Caenorhabditis elegans
CC (see W31268-74) and are considered also to be Wnt receptors. Wnt
CC receptors can be used in a novel, claimed method of screening for
CC compounds which modulate the binding of a Wnt polypeptide (secreted
CC proteins involved in cell-to-cell signalling) to a Wnt receptor.
CC Wnt is involved in (mammary) cancer and other processes involving
CC growth, development and proliferation (both normal and abnormal).
CC Modulators identified by the claimed method are useful for
CC treatment of diseases related to these conditions.
SQ Sequence 585 AA;

Query Match 16.7%; Score 374; DB 27; Length 585;
Best Local Similarity 38.8%; Pred. No. 4.69e-26;
Matches 47; Conservative 29; Mismatches 41; Indels 4; Gaps 4;

Db 20 lvrataaaskapvcqei-t-vpmcrgignylthmpnqfnhdtqdeaglevhgf-pivei 76
QY 27 LFQOPDFSYKRSNCKPIPANLQCHGIEYQNMRLPNLLGHETMKVEY-LEQAGAWIPLVMK 85

Db 77 qcspdlrffictmytpicldpdyhkpccrsvcrakagcplmrgygfawpdmrcdrl 136
QY 86 QCHPDKRFKFLCSLFAFVCVCLDLETTQPCHSILCVQVQKDRCAPVNSAFGFPWMDLECDREP 145

Db 137 p 137
QY 146 P 146

RESULT 4
ID W31270 standard; Protein; 537 AA.
AC W31270;
DE 27-APR-1998 (first entry)
DT Mouse frizzled-4 protein Mfz4 (Wnt receptor).
KW Wnt receptor; mouse frizzled-4 protein; Mfz4 gene;
KW signal transduction; cancer; cell growth; cell proliferation.
OS Mus musculus.
PN W09739357-A1.
PD 23-OCT-1997.
PF 11-APR-1997; U06049.
PR 12-APR-1996; US-015307.
PA (UYJO) UNIV JOHNS HOPKINS.
PA (STRD) UNIV LELAND STANFORD JUNIOR.
PI Andrew D, Bhanot P, Brink M, Hsieh J, Nathans J,
PI Nusse R, Samos CH, Wangy;
DR N-PSDB; T89889.
PT Identification of Wnt receptor binding modulators - useful for
PT treatment of cancer and growth, development or proliferation related
PT disorders
PS Disclosure; Page 34-35; 61pp; English.
CC This protein comprises the mouse transmembrane receptor,
CC frizzled-4 (Mfz4), encoded by the Mfz4 gene (see T89888). It is
CC an example of a Wnt receptor. Other novel frizzled family members
CC have been identified in human, mouse and Caenorhabditis elegans
CC (see W31268-74) and are considered also to be Wnt receptors. Wnt
CC receptors can be used in a novel, claimed method of screening for
CC compounds which modulate the binding of a Wnt polypeptide (secreted
CC proteins involved in cell-to-cell signalling) to a Wnt receptor.
CC Wnt is involved in (mammary) cancer and other processes involving
CC growth, development and proliferation (both normal and abnormal).
CC Modulators identified by the claimed method are useful for

RESULT	11
ID	W35946 standard; Protein; 604 AA.
AC	W35946;
DT	11-MAY-1998 (first entry)
DE	Human netrin-1.
KW	Netrin-1; neuron; growth; differentiation; morphology;
KW	neural disease; diagnosis; therapy; drug screening; human.
OS	Homo sapiens.
PN	WO9740064-A1.

PR 19-APR-1996; US-635137.
PA (EXEL-) EXELIXIS PHARM. INC.
PA (REGC.) UNIV CALIFORNIA.
PI Kennedy T, Leonardo D, Serafini T, Shylan A, Swimmer C,
PI Tessier-Lavigne M, Zhang Y;
DR WPI: 97-535773/49.
DR N-PSDB; T97129.
PT Human netrin-1 protein and related nucleic acids - useful in
PT modulating neuron growth and screening for compounds for diagnosis
PT or treatment of diseases associated with undesirable growth
PS Claim 1; Page 13-15; 22pp: English.
CC This protein comprises human netrin-1 (see W35946), a protein
CC which is involved in neural axon guidance, and which is especially
CC useful in modulating neural axon outgrowth. Its amino acid
CC sequence was deduced from a cDNA clone (see T97129) isolated from a
CC human foetal brain cDNA library. Neuron growth, differentiation or
CC morphology can be altered by contact with netrin (claimed).
CC Isolated netrin-1 can also be used to screen chemical libraries
CC for candidate drugs suitable for diagnosis or treatment of diseases
CC associated with undesirable neural cell growth, by comparing
CC binding to a netrin binding target with and without the presence of
CC a prospective agent. Agents that modulate the interaction may be
CC useful as pharmaceutical lead compounds (claimed).
SQ Sequence 604 AA;

Query Match 5.43; Score 121; DB 28; Length 504;
Best Local Similarity 24.8%; Pred. No. 8.1ee-02;
Matches 29; Conservative 30; Mismatches 49; Indels 9; Gaps

[illegible]

FT	misc_difference	296.7	/note=	"unidentified amino acids"
FT	misc_difference	300.301	/note=	"unidentified amino acids"
FT	misc_difference	307.308	/note=	"unidentified amino acids"
FT	misc_difference	311	/note=	"unidentified amino acid"
FT	misc_difference	389	/note=	"unidentified amino acid"
FT	misc_difference	393.394	/note=	"unidentified amino acids"

```

FT      misc_difference 407
FT      /note= "unidentified amino acid"

```

PN W09513367-A1.
 PD 18-MAY-1995.
 PF 08-NOV-1994; U12913.
 PR 12-NOV-1993; US-152019.
 PA (REGC) UNIV CALIFORNIA.
 PA (UYCO) UNIV COLUMBIA NEW YORK.
 PI Dodd J, Jessell T, Kennedy T, Placzek M, Serafini T;
 PI Tessier-Lavigne M;
 PI WPI: 95-194086/25.
 DR N-PSDB: Q92388.
 PT Neural axon out-growth modulators derived from EGF-like repeats of
 PT netrin 1 or netrin 2 - comprise peptide(s) capable of selectively
 PT increasing spinal axon out-growth or directing axon orientation
 PS Claim 1; Page 51-52; 58pp; English.
 CC Chick p75 and p78 sequences (given in R74186-87, respectively)
 CC were used to identify conserved amino acid regions, which were
 CC then used to design degenerate primers for the amplification
 CC of a fragment of the mouse p78 cDNA (Q92388). This cDNA can
 CC be expressed in host cells for recombinant p78 prodn, or used
 CC to breed transgenic animals, or for gene therapy.
 SQ Sequence 529 AA;
 Query Match 4.7%; Score 105; DB 14; Length 529;
 Best Local Similarity 23.5%; Pred. No. 1.70e+00;
 Matches 24; Conservative 25; Mismatches 45; Indels 8; Gaps 7;
 Db 398 sxveexedcdsckaskgkikmmkkyckrdyavqihilkadgaw-wkftvniis-- 444
 Qy 163 PATEAPKVCBA-CNKNDNDNDMETLCKNDFAKTKV-KEITYINDYKILLETSKT 220
 Db 445 vykqgtsrirtgds-lwirsrdackepkikplk-ylllg 484
 Qy 221 IYKLVGVSERLKKSVLWKD-SLQCTCEMNDINAPLYVMG 261
 RESULT 13
 ID W26642 standard; Protein; 732 AA.
 AC W26642;
 DT 11-FEB-1998 (first entry)
 DE Human RECK cancer-inhibiting protein.
 KW RECK; reversion-inducing cysteine rich protein with Kazal motif;
 KW human; cancer; gene therapy; diagnosis.
 OS Homo sapiens.
 FH Key Location/Qualifiers
 FT Domain 396..415
 FT /label= Kazal_domain
 PN W09724439-A1.
 PD 10-JUL-1997.
 PF 24-DEC-1996; U20812.
 PR 27-DEC-1995; JP-340469.
 PA (AMGE-) AMGEN INC.
 PA (KITA/) KITAYAMA H.
 PA (NODA/) NODA M.
 PA (TAKA/) TAKAHASHI C.
 PA (SANY) SANKYO CO LTD.
 PI Kitayama H, Noda M, Takahashi C;
 PI WPI: 97-363675/33.
 DR N-PSDB: T90508.
 PT RECK gene and corresponding protein sequences - enables reversion of
 PT cancer cells
 PS Claim 7; Page 42-44; 53pp; English.
 CC This protein sequence comprises the human RECK protein (reversion-
 CC inducing cysteine rich protein with Kazal motif), which is capable
 CC of transforming malignant cancer cells with an activated ras gene
 CC into normal cells (reversion activity). Its amino acid sequence
 CC was deduced from a cDNA clone (see T90508) obtained from human
 CC fibroblast MRC-5 (ATCC CCL 171) cells. A claimed method for
 CC treatment of cancer comprises contacting the cancer cells with a
 CC RECK polypeptide. RECK may also be expressed using gene therapy
 CC methods for in vivo treatment of cancer.
 SQ Sequence 732 AA;
 Query Match 4.5%; Score 100; DB 25; Length 732;

Best Local Similarity 30.5%; Pred. No. 4.26e+00;
 Matches 25; Conservative 15; Mismatches 36; Indels 6; Gaps 5;
 Db 158 ipvldikkcpemwkaiaclqikpchkshksrsii-cksdceveillkkcgd-qnkfpdht 215
 Qy 80 IP-LVMKQCHPDT-KKFLCSLFAFVCLDLDETQPCSLCVQVKRCAPVMSAFGFPWP 137
 Db 216 aesicellsptddkncipldt 237
 Qy 138 DMLECDRFPQDNDL--CIPLAS 157
 RESULT 14
 ID W55368 standard; Protein; 804 AA.
 AC W55368;
 DT 17-JUN-1998 (first entry)
 DE H. pylori ORF 09apl1406orf2 protein.
 KW Cytoplasmic; vaccine; prevention; treatment; infection; envelope;
 KW identification; binding compound; bacteria; life cycle; activator;
 KW inhibitor; duodenal ulcer disease; chronic gastritis; diagnosis.
 OS Helicobacter pylori.
 PN W09737044-A1.
 PD 09-OCT-1997.
 PF 27-MAR-1997; U05223.
 PR 06-DEC-1996; US-761318.
 PR 29-MAR-1996; US-625811.
 PR 02-APR-1996; US-758731.
 PR 25-OCT-1996; US-736905.
 PR 28-OCT-1996; US-738859.
 PA (ASTR) ASTRA AB.
 PI Alm RA, Smith D;
 PI WPI: 97-503122/46.
 DR N-PSDB: V24777.
 PT Helicobacter pylori nucleic acid sequences and encoded
 PT polypeptide(s) - useful in vaccines to treat or prevent H. pylori
 PT infection and for diagnosis of H. pylori infection
 PS Claim 14; Page 587-589; 1145pp; English.
 CC This sequence is a H. pylori protein of unspecified function.
 CC The protein may be used in a vaccine to prevent or treat H. pylori
 CC infection or to identify H. pylori polypeptide binding compounds. The
 CC useful as potential H. pylori life cycle activators or inhibitors. The
 CC DNA and probes derived from it may be used for the identification of
 CC H. pylori in a sample and the diagnosis of H. pylori infection. Nucleic
 CC acid sequences complementary to the DNA act as antisense sequences and
 CC can be used to prevent the translation of H. pylori mRNA. Antibodies
 CC against the protein can be used in immunoassays to evaluate the abundance
 CC and distribution of H. pylori-specific antigens. The genomic sequence of
 CC H. pylori (ATCC 35679) was determined from overlapping contigs generated
 CC by mechanically shearing the bacterial DNA. The sequences were analysed
 CC for ORF of at least 180 nucleotides, and the predicted coding regions
 CC defined by computer evaluation. To identify likely H. pylori antigens for
 CC vaccine development, the amino acid sequences predicted from various ORF
 CC were analysed for significant homology to other known or exported
 CC membrane proteins. Having identified and determined the sequences of
 CC interest, particular regions can be isolated from H. pylori by PCR
 CC amplification for recombinant polypeptide production, e.g. in E. coli
 CC hosts.
 SQ Sequence 804 AA;
 Query Match 4.4%; Score 99; DB 29; Length 804;
 Best Local Similarity 31.5%; Pred. No. 5.11e+00;
 Matches 17; Conservative 16; Mismatches 19; Indels 2; Gaps 2;
 Db 78 vidsipkegftplenaftnysmrqgqfihikpkmvrltlfdfdrdykai 131
 Qy 185 IMETLCKNDFA-LKIKVKEITINRDTKILLETKSKTIYKLVGS-ERDLKKS 236
 RESULT 15
 ID W26610 standard; Protein; 868 AA.
 AC W26610;
 DT 27-JAN-1998 (first entry)
 DE Rat muscle-specific kinase (MUSK).

Search completed: Tue Aug 4 10:32:08 1998
Job time : 48 secs.

THIS PAGE BLANK (USPTO)

WIREH

(TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

MPSrch_pp protein - protein database search, using Smith-Waterman algorithm
Run on: Tue Aug 4 10:32:25 1998; MasPar time 16.16 Seconds
Tabular output not generated. 667.011 Million cell updates/sec

Title: >US-08-848-439-2
Description: (1-295) from US08848439.pep
Perfect Score: 2234
Sequence: 1 MLQPGSLLLLFLASHCCLG.....WKGOREFKRISIRKLCQ 295

Scoring table: PAM 150
Gap 11

Searched: 120441 seqs, 36531193 residues

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: pir56
1:pir1 2:pir2 3:pir3 4:pir4 5:nrl3d

Statistics: Mean 44.927; Variance 86.902; scale 0.517

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	386	17.3	641	2	A45054	1.70e-52
2	355	15.9	694	2	S78444	2.44e-46
3	355	15.9	694	2	S71786	2.44e-46
4	332	14.9	581	2	S15709	8.11e-42
5	332	14.9	581	2	S03540	8.11e-42
6	178	8.0	562	2	B56101	3.36e-13
7	148	6.6	581	2	B54665	3.23e-08
8	124	5.6	606	2	A54655	1.67e-04
9	122	5.5	793	2	JC5359	3.30e-04
10	105	4.7	500	2	D31579	8.40e-02
11	105	4.7	500	2	I52313	8.40e-02
12	101	4.5	1735	2	S22812	2.87e-01
13	99	4.4	322	2	H54825	5.25e-01
14	98	4.4	391	2	E54474	7.08e-01
15	98	4.4	724	2	B32571	5.25e-01
16	98	4.4	1032	1	G7BP74	7.08e-01
17	96	4.3	520	2	J50291	1.28e+00
18	96	4.3	735	2	I51901	1.28e+00
19	97	4.3	15281	2	S41309	9.52e-01
20	94	4.2	271	2	E54694	2.28e+00
21	93	4.2	735	2	A53300	3.04e+00
22	92	4.1	203	2	B70254	4.05e+00
23	92	4.1	608	2	B70188	4.05e+00

24	92	4.1	801	2	B64579	paralysed flagella pr	4.05e+00
25	92	4.1	1562	2	S53069	probable membrane pro	4.05e+00
26	90	4.0	176	2	S70008	tropoin I - Atlantic	7.11e+00
27	89	4.0	262	2	A60105	virf protein - Shigel	9.38e+00
28	89	4.0	262	2	A47605	virf virulence protei	9.38e+00
29	89	4.0	262	2	S14846	virf protein - Shigel	9.38e+00
30	90	4.0	283	2	S73489	S-adenosylmethionine-	7.11e+00
31	89	4.0	508	2	A64571	GMP synthase - Helico	9.38e+00
32	90	4.0	521	2	A53153	glucose transport pro	7.11e+00
33	89	4.0	752	2	A32571	ribosomal protein S6	9.38e+00
34	88	3.9	137	2	S45967	probable membrane pro	1.24e+01
35	88	3.9	182	2	S28310	hypothetical protein	1.24e+01
36	88	3.9	183	2	D69108	phycoerythrin alpha ph	1.24e+01
37	88	3.9	370	2	A69430	conserved hypothetical	1.24e+01
38	88	3.9	443	2	JQ1527	alpha-amylase (EC 3.2	1.24e+01
39	88	3.9	445	2	S19990	alpha-amylase (EC 3.2	1.24e+01
40	88	3.9	633	2	C32571	ribosomal protein S6	1.24e+01
41	88	3.9	740	2	I38556	insulin-stimulated pr	1.24e+01
42	88	3.9	849	2	S00030	neurofilament triplet	1.24e+01
43	88	3.9	871	2	I48696	gene Nsk2 protein - m	1.24e+01
44	88	3.9	881	2	I48697	gene Nsk2 protein - m	1.24e+01
45	88	3.9	1102	2	S44772	C2954.4 protein - Cae	1.24e+01

ALIGNMENTS

RESULT 1 A45054 #type complete
ENTRY probable intercellular signal transducer or transmitter Fz-1
TITLE - rat
ORGANISM #formal_name Rattus norvegicus #common_name Norway rat
DATE 30-Apr-1993 #sequence_revision 18-Nov-1994 #text_change 28-Apr-1995
A45054
A45054
Chan, S.D.; Karpf, D.B.; Fowlkes, M.E.; Hooks, M.; Bradley, M.S.; Vuong, V.; Bambino, T.; Liu, M.Y.; Arnaud, C.D.; Strewler, G.J.; Nissenson, R.A.
J. Biol. Chem. (1992) 267:25202-25207
Two homologs of the Drosophila polarity gene frizzled (Fz) are widely expressed in mammalian tissues.
#cross-references MUID:93094228
#accession A45054
#status preliminary; not compared with conceptual translation
#molecule_type nucleic acid
#residues 1-641 #label CHA
#experimental_source UMR 106 osteosarcoma cell line
#note sequence extracted from NCBI backbone (NCBIP:120154) #length 641 #molecular_weight 71054 #checksum 8376

Query Match 17.3%; Score 386; DB 2; Length 641;
Best Local Similarity 39.2%; Pred. No. 1.70e-52;
Matches 49; Conservative 27; Mismatches 43; Indels 6; Gaps 6;
Db 108 HGYPQI-S-IPICTDIAYNQITMPNLLGHTNOEDAGLEVHQYPLVKVOCSELKFFLC 165
QY 37 RSNCKIPANLQCHGIEYQNMELPNLLGHETMKVELEQAGAPIPLVMKQCHPDTRKFLC 96
Db 166 SMYAPVC-TVLEQALPPCSLCERAG-CFALNKKFGFWPDTLKCEKFPVHGELCVG 223
QY 97 SLFAPVCLDDLETIQCHSLCVQVQKDRCAPVMSAFGPMPLMCDRFP-QD-NDLCIP 154
Db 224 QNTSD 228
QY 155 LASSD 159

RESULT 2 S78444 #type complete
ENTRY dfz2 protein - fruit fly (Drosophila melanogaster)
TITLE #formal_name Drosophila melanogaster
ORGANISM 12-Feb-1998 #sequence_revision 13-Mar-1998 #text_change 13-Mar-1998
DATE

```
ACCESSIONS S78444
REFERENCE S78444
#authors Bhanot, P.; Wang, Y.; Nathans, J.
#submission submitted to the EMBL Data Library, July 1996
#accession S78444
##status preliminary
##molecule_type DNA
##residues 1-694 #label BHA
##cross-references EMBL:U65589
#length 694 #molecular-weight 75437 #checksum 3113

SUMMARY
Query Match 15.9%; Score 355; DB 2; Length 694;
Best Local Similarity 36.2%; Pred. No. 2,44e-46;
Matches 47; Conservative 32; Mismatches 45; Indels 6; Gaps 6;

Db 51 YGVPAIP-KDPNLRCEIIPMCRGIGYNTSFPNEMNHETODEAGLEVHQFW-PLVEIK 108
QY 28 FGQPDFSYKRSNCKPIPANLQCHGIEYQNMRLPNLLGHETMKEV-LEQAGAWIPLVMKQ 86
Db 109 CSPDLKFFLCMSYTPICLDYHKPLPVCRSVCERARSGCAPIMQOYSFEWPERMACEHLP 168
QY 87 CHPDTKKFLCSLFAPVCLDDLETIQPCHSLCVQVKDRCAPVMSAFGFPWPDMLECDREP 146
Db 169 LHGDPDNLCLM 178
QY 147 -Q-D-NDLCI 153

RESULT 3
ENTRY S71786 #type complete
TITLE dfz2 protein - fruit fly (Drosophila melanogaster)
ORGANISM #formal_name Drosophila melanogaster
DATE 12-Feb-1998 #sequence_revision 13-Mar-1998 #text_change
13-Mar-1998
ACCESSIONS S71786
REFERENCE S71786
#authors Bhanot, P.; Brink, M.; Harryman Samos, C.; Hsieh, J.C.; Wang,
Y.; Macke, J.P.; Andrew, D.; Nathans, J.; Nusse, R.
#journal Nature (1996) 382:225-230
#title A new member of the frizzled family from Drosophila functions
as a wingless receptor.
#accession S71786
##status preliminary; nucleic acid sequence not shown
##molecule_type DNA
##residues 1-694 #label BHA
##cross-references EMBL:U65589
#length 694 #molecular-weight 75423 #checksum 3095

SUMMARY
Query Match 15.9%; Score 355; DB 2; Length 694;
Best Local Similarity 36.2%; Pred. No. 2,44e-46;
Matches 47; Conservative 32; Mismatches 45; Indels 6; Gaps 6;

Db 51 YGVPAIP-KDPNLRCEIIPMCRGIGYNTSFPNEMNHETODEAGLEVHQFW-PLVEIK 108
QY 28 FGQPDFSYKRSNCKPIPANLQCHGIEYQNMRLPNLLGHETMKEV-LEQAGAWIPLVMKQ 86
Db 109 CSPDLKFFLCMSYTPICLDYHKPLPVCRSVCERARSGCAPIMQOYSFEWPERMACEHLP 168
QY 87 CHPDTKKFLCSLFAPVCLDDLETIQPCHSLCVQVKDRCAPVMSAFGFPWPDMLECDREP 146
Db 169 LHGDPDNLCLM 178
QY 147 -Q-D-NDLCI 153

RESULT 4
ENTRY S15709 #type complete
TITLE hypothetical protein - fruit fly (Drosophila melanogaster)
ORGANISM #formal_name Drosophila melanogaster
DATE 21-Nov-1993 #sequence_revision 10-Nov-1995 #text_change
16-Feb-1997
ACCESSIONS S15709
REFERENCE S15708
```

```
#authors Adler, P.N.; Vinson, C.; Park, W.J.; Conover, S.; Klein, L.
#journal Genetics (1990) 126:401-416
#title Molecular structure of frizzled, a Drosophila tissue polarity
gene.
#cross-references MUID:91060073
#accession S15709
##status preliminary
##molecule_type DNA
##residues 1-415 #label ADL
##cross-references EMBL:X54648
GENETICS
#gene FlyBase: fz
##cross-references FlyBase: FBgn0001085
#introns 224/3; 264/3; 329/3; 405/1
SUMMARY #length 415 #molecular-weight 46117 #checksum 4401

Query Match 14.9%; Score 332; DB 2; Length 415;
Best Local Similarity 37.5%; Pred. No. 8,11e-42;
Matches 45; Conservative 22; Mismatches 47; Indels 6; Gaps 6;

Db 50 HNRCEPI-T-ISICKNIPYNTIMPNLIGHTKQEEAGLEVHQFAPLVKIGSCDDLQFLC 107
QY 37 RSNCAPIFANLQCHGIEYQNMRLPNLLGHETMKEVLEQAGAWIPLVMKQCHPDTKKFLC 96
Db 108 SLYVPVC-TILERPPIPCRSLSARV-CEKLMKTYNFWPENLECSKFPVHGGEDLCVA 165
QY 97 SLFAPVCLDDLETIQPCHSLCVQVKDRCAPVMSAFGFPWPDMLECDREP-QD-NDLCIP 154

RESULT 5
ENTRY S03540 #type complete
TITLE gene frizzled protein precursor - fruit fly (Drosophila
melanogaster)
ORGANISM #formal_name Drosophila melanogaster
DATE 31-Mar-1990 #sequence_revision 31-Mar-1990 #text_change
23-Aug-1997
ACCESSIONS S03540; S15708
REFERENCE S03540
#authors Vinson, C.R.; Conover, S.; Adler, P.N.
#journal Nature (1989) 338:263-264
#title A Drosophila tissue polarity locus encodes a protein
containing seven potential transmembrane domains.
#cross-references MUID:89159415
#accession S03540
##status not compared with conceptual translation
##molecule_type mRNA
##residues 1-581 #label VIN
REFERENCE S15708
#authors Adler, P.N.; Vinson, C.; Park, W.J.; Conover, S.; Klein, L.
#journal Genetics (1990) 126:401-416
#title Molecular structure of frizzled, a Drosophila tissue polarity
gene.
#cross-references MUID:91060073
#accession S15708
##status preliminary
##molecule_type DNA
##residues 1-581 #label ADL
##cross-references EMBL:X54648
GENETICS
#gene FlyBase: fz
##cross-references FlyBase: FBgn0001085
#introns 224/3; 264/3; 329/3; 405/1
KEYWORDS alternative splicing; transmembrane protein
FEATURE 1-26 #domain signal sequence #status predicted #label SIG\
27-581 #product gene frizzled protein #status predicted #label
MAT
SUMMARY #length 581 #molecular-weight 64846 #checksum 358

Query Match 14.9%; Score 332; DB 2; Length 581;
Best Local Similarity 37.5%; Pred. No. 8,11e-42;
Matches 45; Conservative 22; Mismatches 47; Indels 6; Gaps 6;
```

```

Db 50 HNRCPPI-T-ISICKNIPYNWTIMPNLIGHTKQEPAGLEVHQFAPLVYKIGSGDDDLQLFLC 107
::||::||::||::||::||::||::||::||::||::||::||::||::||::||::||::||::||
Qy 37 RSNCKPIPANLQCHGIEYQNMRLPNLLGHETMKVELEQAGAWIPLVYKQCHPDTKKFLC 96
::||::||::||::||::||::||::||::||::||::||::||::||::||::||::||::||::||
Db 108 SLYPVC-TILERTPPCRSICSAHV-CEKLMKTYNWPENLECSKFPVHGGEGLCA 165
||::||::||::||::||::||::||::||::||::||::||::||::||::||::||::||::||
Qy 97 SLFAPVCLDDLDLDTIQPCHSLCQVQKDRCAVPMSAFGFPWPMLECDREFP-QD-NDLCIP 154
||::||::||::||::||::||::||::||::||::||::||::||::||::||::||::||::||

RESULT 6
ENTRY B56101 #type fragment
TITLE collagen alpha 1(XVII) chain precursor long form - mouse
ORGANISM #formal_name Mus musculus #common_name house mouse
DATE 03-Oct-1995 #sequence_revision 03-Oct-1995 #text_change
10-Sep-1997
ACCESSIONS B56101
REFERENCE A56101
#authors Rehn, M.; Pihlajaniemi, T.
#journal J. Biol. Chem. (1995) 270:4705-4711
#title Identification of three N-terminal ends of type XVIII
collagen chains and tissue-specific differences in the
expression of the corresponding transcripts. The longest
form contains a novel motif homologous to rat and
Drosophila frizzled proteins.
#accession B56101
##status preliminary
##molecule_type mRNA
##residues 1-562 #label REH
##cross-references GB:U11637; NID:9618429; PID:9618430
GENETICS
#gene COL18A1
SUMMARY #length 562 #checksum 115

Query Match 8.0%; Score 178; DB 2; Length 562;
Best Local Similarity 34.1%; Pred. No. 3.36e-13;
Matches 30; Conservative 13; Mismatches 45; Indels 0; Gaps 0;

Db 368 SRCLPPLPTLCRLGIGFWPLNHLHTDSVEATVQAWGRFLHTNCPFLAWFFCL 427
||::||::||::||::||::||::||::||::||::||::||::||::||::||::||::||::||
Qy 38 SNCRPIPANLQCHGIEYQNMRLPNLLGHETMKVELEQAGAWIPLVYKQCHPDTKKFLC 97
||::||::||::||::||::||::||::||::||::||::||::||::||::||::||::||::||

Db 428 LLAPSCGPPPPPLPPCRQCFCEALEDEC 455
||::||::||::||::||::||::||::||::||::||::||::||::||::||::||::||::||
Qy 98 LFAPVCLDDLDLDTIQPCHSLCQVQKDRC 125
||::||::||::||::||::||::||::||::||::||::||::||::||::||::||::||::||

RESULT 7
ENTRY B54665 #type fragment
TITLE netrin-2 precursor - chicken (fragment)
ORGANISM #formal_name Gallus gallus #common_name chicken
DATE 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change
10-Sep-1997
ACCESSIONS B54665
REFERENCE A54665
#authors Serafini, T.; Kennedy, T.E.; Galko, M.J.; Mirzayan, C.;
Jessel, T.M.; Tessier-Lavigne, M.
#journal Cell (1994) 78:409-424
#title The netrins define a family of axon outgrowth-promoting
proteins homologous to Caenorhabditis elegans UNC-6.
#accession B54665
##status preliminary
##molecule_type mRNA
##residues 1-581 #label SER
##cross-references GB:L34550; NID:9529420; PID:9529421
SUMMARY #length 581 #checksum 7731

Query Match 6.6%; Score 148; DB 2; Length 581;
Best Local Similarity 27.0%; Pred. No. 3.23e-08;
Matches 30; Conservative 28; Mismatches 46; Indels 7; Gaps 6;

Db 431 IPAINPTSLVSTSTE-APADCSYCKPAKGNKYKINMKYCKDKYVQVNVNILEMETVANWAK 489
||::||::||::||::||::||::||::||::||::||::||::||::||::||::||::||::||

```

```

Qy 153 IPLASSDHLLPATEAPKVCEA-CKKNDDDDNDIMETLCKNDNFALKIKVKEITYINRDTK 211
::||::||::||::||::||::||::||::||::||::||::||::||::||::||::||::||
Db 490 FTNILS--VYKCDERVKR-GDNFLMHLKDLSCCKPKI-QISKKYLVMG 536
::||::||::||::||::||::||::||::||::||::||::||::||::||::||::||::||
Qy 212 IILETKSKTIYKLVGVSERDLKSLVWLK-DSLOCTCEENNDINAPVLMG 261
||::||::||::||::||::||::||::||::||::||::||::||::||::||::||::||::||

RESULT 8
ENTRY A54665 #type complete
TITLE netrin-1 precursor - chicken
ORGANISM #formal_name Gallus gallus #common_name chicken
DATE 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change
10-Sep-1997
ACCESSIONS A54665
REFERENCE A54665
#authors Serafini, T.; Kennedy, T.E.; Galko, M.J.; Mirzayan, C.;
Jessel, T.M.; Tessier-Lavigne, M.
#journal Cell (1994) 78:409-424
#title The netrins define a family of axon outgrowth-promoting
proteins homologous to Caenorhabditis elegans UNC-6.
#accession A54665
##status preliminary
##molecule_type mRNA
##residues 1-606 #label SER
##cross-references GB:L34549; NID:9529418; PID:9529419
SUMMARY #length 606 #molecular-weight 68126 #checksum 5400

Query Match 5.6%; Score 124; DB 2; Length 606;
Best Local Similarity 26.1%; Pred. No. 1.67e-04;
Matches 31; Conservative 34; Mismatches 41; Indels 13; Gaps 11;

Db 456 IPAAPPTTASSTEEPAD-CDSYCKASKGKLKINMKYCKDYAVQIHLKAENADM-W 513
||::||::||::||::||::||::||::||::||::||::||::||::||::||::||::||::||
Qy 153 IPLASSDHLLPATEAPKVCEA-CKKNDDDDNDIMETLCKNDNFALKIKV-KEITYINRDT 210
||::||::||::||::||::||::||::||::||::||::||::||::||::||::||::||::||

Db 514 KFTVNIIS--VYK-OG-SNR-LRRGDOTLVHAKDIACKCPKVPMKK-YLLLGSTEDS 566
||::||::||::||::||::||::||::||::||::||::||::||::||::||::||::||::||
Qy 211 KIILETKSKTIYKLVGVSERDLKKS--VLWLKD-SLOCTCEENNDINAPVLMGQKGG 266
||::||::||::||::||::||::||::||::||::||::||::||::||::||::||::||::||

RESULT 9
ENTRY JC5539 #type complete
TITLE Smoothed protein - mouse
ORGANISM #formal_name Mus musculus #common_name house mouse
DATE 02-Sep-1997 #sequence_revision 05-Sep-1997 #text_change
05-Sep-1997
ACCESSIONS JC5539; PC4476
REFERENCE JC5539
#authors Akiyama, H.; Shigeno, C.; Hiraki, Y.; Shukunami, C.; Kohnno,
H.; Akagi, M.; Konishi, J.; Nakamura, T.
#journal Biochem. Biophys. Res. Commun. (1997) 235:142-147
#title Cloning of a mouse Smoothed cDNA and expression patterns of
hedgehog signalling molecules during chondrogenesis and
cartilage differentiation in clonal mouse EC cells, ADPC5.
#accession JC5539
##molecule_type mRNA
##residues 1-793 #label AKI
#accession PC4476
##molecule_type protein
##residues 528-533:539-545; 600-605 #label AK2
##experimental_source ADPC5 cell
COMMENT This protein is used in the conserved targets in Hedgehog
signalling pathway, together with Patched and Gli. These protein
are responsible for the skeletal abnormalities in Gorlin and
Greig syndromes.
FEATURE
1-32 #domain signal sequence #status predicted #label SIG
SUMMARY #length 793 #molecular-weight 87299 #checksum 8609

Query Match 5.5%; Score 122; DB 2; Length 793;
Best Local Similarity 26.7%; Pred. No. 3.30e-04;
Matches 31; Conservative 22; Mismatches 54; Indels 9; Gaps 7;

```

```

ACCESSIONS      152313
REFERENCE
#authors        Kawashima, H.; Strobel, H.W.
#journal        Biochem. Biophys. Res. Commun. (1995) 209:535-540
#title          cDNA cloning of a novel rat brain cytochrome P450 belonging
                to the CYP2D subfamily.
#cross-references MUID:95251650
#accession      152313
##status        preliminary; translated from GB/EMBL/DBJ
##molecule_type mRNA
##residues       1-500 #label RES
##cross-references GB:S77859; NID:g998524; PID:g998525
##experimental_source brain, strain Sprague-Dawley
CLASSIFICATION  #superfamily cytochrome P450
KEYWORDS         heme; transmembrane protein
FEATURE
FEATURE_446     #binding_site heme iron (Cys) (axial ligand) #status
                predicted
SUMMARY          #length 500 #molecular-weight 56603 #checksum 8540
                Predicted
Query Match      4.7%; Score 105; DB 2; Length 500;
Best Local Similarity 24.7%; Pred.No. 8.40e-02;
Matches 18; Conservative 28; Mismatches 21; Indels 6; Gaps 5

Db   160 ARCICAAPADHSGPPFS--PNTLLDKAVCNVIASSLFACR-FEYNDFRIRLL--DLLKDT 215
      :|:: :::: ::||:: :::: :::: :::: :::: :::: :::: :::: :::: ::::
QY   15 SHC-CLG-SARGLFLFGQPDSFYKRSKCKPIPANQLCHGIEYQNMLPLNLIGHETMKEV 72
      ::|| :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::

Db   216 LSEESGFPLMLLN 228
      ||:: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::
QY   73 LEQAGAWIPLYNK 85

RESULT 12
ENTRY   S22812           #type complete
TITLE   DNA-directed RNA polymerase (EC 2.7.7.6) III largest chain
ORGANISM Giardia lamblia
DATE    12-Feb-1993 #sequence_revision 12-Feb-1993 #text_change
        12-Sep-1997
ACCESSIONS      S22812
REFERENCE       S22811
#authors        Lanzenodoerfer, M.; Palm, P.; Grampp, B.; Peattie, D.A.;
                Zillig, W.
#journal        Nucleic Acids Res. (1992) 20:1145
#title          Nucleotide sequence of the gene encoding the largest subunit
                Of the DNA-dependent RNA polymerase III of Giardia lamblia
#cross-references MUID:92195823
#accession      S22812
##status        nucleic acid sequence not shown; translation not shown
##molecule_type DNA
##residues       1-1735 #label LAN
##cross-references EMBL:X60325
##note          the nucleotide sequence was submitted to the EMBL Data
                Library, September 1991

GENETICS
#gene           ipoA3
CLASSIFICATION  #superfamily Giardia DNA-directed RNA polymerase III larges
Chain
KEYWORDS         DNA binding; nucleotidyltransferase; nucleus; transcription
                zinc finger
SUMMARY          #length 1735 #molecular-weight 193191 #checksum 769
                4.5%; Score 101; DB 2; Length 1735;
Query Match      27.4%; Pred.No. 2.87e-01;
Best Local Similarity 20; Conservative 25; Mismatches 23; Indels 5; Gaps 5

Db   168 TICKKSQGRV-LRMHYVEERATKILADVTKHKQTEKIQLDQR-LRNHVVDPIQALHI 225
      |::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::
QY   188 TlCKNDfALKIKvEIYINRDtKIILET-KS-KTIYKLNGVSERDLKKSVLWLDKSLQC 245
      :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::

Db   226 L-QKVPECDIPYL 237
      :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::

```

[illegible]

THIS PAGE BLANK (USPTO)

W E I R D

(TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

Mpsrch_pp protein - protein database search, using Smith-Waterman algorithm
Run on: Tue Aug 4 10:33:44 1998; Maspar time 10.67 Seconds
Tabular output not generated. 693.479 Million cell updates/sec

Title: >US-08-848-439-2
Description: (1-295) from US08848439.pep
Perfect Score: 2234
Sequence: 1 MLQPGSGLLLLFLASHCCLG.....WQKGQREKFRKRSIRKLQC 295

Scoring table: PAM 150
Gap 11

Searched: 69111 seqs, 25083644 residues

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: swiss-prot35
1:swiss1

Statistics: Mean 46.742; Variance 71.562; scale 0.653

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	332	14.9	581	1	FRIZ_DROME	9.66e-53
2	148	6.6	581	1	NETRIN-2 PRECURSOR (FR	8.14e-11
3	124	5.6	606	1	NETRIN-1 PRECURSOR.	3.06e-06
4	109	4.9	465	1	YQ18_CAEEL	1.32e-03
5	105	4.7	500	1	CYP4_RAT	6.13e-03
6	105	4.7	500	1	CPD1_RAT	2.73e-02
7	101	4.5	1741	1	PCP1_GIALA	6.13e-02
8	99	4.4	724	1	YQ18_MOUSE	5.67e-02
9	98	4.4	1032	1	BASEPLATE STRUCTURAL P	8.13e-02
10	96	4.3	520	1	ION3_CARAU	1.66e-01
11	94	4.2	531	1	UDP-GLUCURONOSYLTRANSF	3.35e-01
12	92	4.1	1562	1	YQ18_YEAST	6.69e-01
13	90	4.0	216	1	GYRB_ACIS3	1.31e+00
14	89	4.0	262	1	VIRF_SHIDY	1.83e+00
15	90	4.0	263	1	KSGA_MFCPN	1.31e+00
16	90	4.0	561	1	VATA_MAIZE	1.83e+00
17	89	4.0	752	1	KSGA_CHICK	1.83e+00
18	88	3.9	127	1	YB9_YEAST	2.55e+00
19	88	3.9	182	1	YLP2_CAEEL	2.55e+00
20	88	3.9	443	1	AMZA_ORISA	2.55e+00
21	88	3.9	445	1	AMC2_ORISA	2.55e+00
22	87	3.9	446	1	YHE2_PSEAE	3.53e+00
23	87	3.9	510	1	NOA1_HUMAN	3.53e+00

24	87	3.9	580	1	VATA_HORVU	3.53e+00
25	88	3.9	633	1	KS62_MOUSE	2.55e+00
26	88	3.9	740	1	KS62_HUMAN	2.55e+00
27	88	3.9	848	1	NEF_MOUSE	2.55e+00
28	88	3.9	1102	1	YK64_CAEEL	2.55e+00
29	87	3.9	1201	1	COPA_YEAST	3.53e+00
30	87	3.9	1752	1	DESP_HUMAN	3.53e+00
31	87	3.9	2386	1	FINC_HUMAN	3.53e+00
32	87	3.9	2749	1	IP3R_MOUSE	3.53e+00
33	87	3.9	2749	1	IP3R_RAT	3.53e+00
34	86	3.8	327	1	VMP_CAMVD	4.88e+00
35	86	3.8	450	1	DCOR_CHICK	4.88e+00
36	86	3.8	461	1	DCOR_BOVIN	4.88e+00
37	85	3.8	495	1	EIBL_ADE02	6.72e+00
38	86	3.8	556	1	FTHS_STRMU	4.88e+00
39	86	3.8	620	1	Y870_METJA	4.88e+00
40	86	3.8	725	1	VACB_MYCCE	4.88e+00
41	85	3.8	798	1	VP16_YEAST	6.72e+00
42	85	3.8	1701	1	MSPI_PLAFC	6.72e+00
43	85	3.8	1726	1	MSPI_PLAFC	6.72e+00
44	85	3.8	1726	1	MSPI_PLAFC	6.72e+00
45	85	3.8	2688	1	2EP1_MOUSE	6.72e+00

ALIGNMENTS

RESULT 1
ID FRIZ_DROME STANDARD; PRT; 581 AA.
AC P18537;
DT 01-NOV-1990 (REL. 16, CREATED)
DT 01-NOV-1990 (REL. 16, LAST SEQUENCE UPDATE)
DT 01-MAY-1992 (REL. 22, LAST ANNOTATION UPDATE)
DE FRIZZLED PROTEIN PRECURSOR.
GN FZ;
OS DROSOPHILA MELANOGASTER (FRUIT FLY).
OC EUKARYOTA; METAZOA; ARTHROPODA; INSECTA; DIPTERA.
RN [1]
RP SEQUENCE FROM N.A. (CLASS I).
RX MEDLINE; 89159415.
RA VINSON C.R., CONOVER S., ADLER P.N.;
RL NATURE 338:263-264(1989).
RN [2]
RP SEQUENCE FROM N.A. (CLASSES I AND II).
RX MEDLINE; 91060073.
RA ADLER P.N., VINSON C., PARK W.J., CONOVER S., KLEIN L.;
RL GENETICS 126:401-416(1990).
CC -!- FUNCTION: REQUIRED TO COORDINATE THE CYTOSKELETONS OF EPIDERMAL CELLS TO PRODUCE A PARALLEL ARRAY OF CUTICULAR HAIRS AND BRISTLES.
CC FZ IS REQUIRED FOR THE TRANSMISSION OF POLARITY INFORMATION, AND TRANSDUCTION.
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC -!- ALTERNATIVE PRODUCTS: CLASS I AND CLASS II PROTEINS DIFFER AT THE C-TERMINAL FROM AA 406 DUE TO USE OF ALTERNATIVE 3' EXONS.
DR EMBL; X54648; G804979; JOINED.
DR EMBL; X54649; G804979; JOINED.
DR EMBL; X54650; G804979; JOINED.
DR EMBL; X54651; G804979; JOINED.
DR EMBL; X54648; G804980; JOINED.
DR EMBL; X54649; G804980; JOINED.
DR EMBL; X54650; G804980; JOINED.
DR EMBL; X54652; G804980; JOINED.
DR EMBL; X54646; G7981; -.
DR EMBL; X54647; G7983; ALT_SEQ.
DR PIR; S03540; S03540.
DR PIR; S15708; S15708.
DR FLYBASE; FBgn0001085; fz.
KW TRANSMEMBRANE; SIGNAL; ALTERNATIVE SPLICING; DEVELOPMENTAL PROTEIN.
FT SIGNAL 1 26
FT CHAIN 27 581
FT FRIZZLED PROTEIN.
FT TRANSMEM 248 270
FT TRANSMEM 281 303
FT TRANSMEM 336 368
FT TRANSMEM 381 401

FT	DISULFID	453	521	BY SIMILARITY.
FT	DISULFID	468	578	BY SIMILARITY.
FT	CARBOHYD	88	88	POTENTIAL.
FT	CARBOHYD	103	103	POTENTIAL.
FT	CARBOHYD	394	394	POTENTIAL.
FT	CARBOHYD	540	540	POTENTIAL.
SEQ	SEQUENCE	581 AA;	65106 MW; 5B6D2272 CRC32;	
Query Match				
Best Local Similarity		27.08;	Score 148; DB 1; Length 581;	
Matches		30; Conservative	28; Mismatches 46; Indels 7; Gaps	
Db	431	IPAINPTSLVTSTE-APADCD	SYCKPAKGYKINMKYKCKDYYVQVNI	LEMETVANWAK 489
Qy	153	IPLASSDHLLPATEEAPK	VCSEA-CKNKDDDDIMETLCKNDFALKIKVKEITY	INRDTK 211
Db	490	FTNILS-VYKCRDERVKR-GDNFL	WIHLKDLSCPKPI-QISKYLYWG	536
Qy	212	IILETKRTIKYKLVGSERDLK	KKSVLWK-DSLOQTCCEMNDINAPYLVG	261
RESULT	3			
ID	NET1_CHICK	STANDARD;	PRT;	606 AA.
AC	Q90922;			
DT	01-NOV-1997 (REL. 35, CREATED)			
DT	01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)			
DT	01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)			
DE	NETRIN-1 PRECURSOR.			
OS	GALLUS GALLUS (CHICKEN).			
OC	EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; AVES; NEOGNATHAE			
OC	GALLIFORMES.			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN-WHITE LEGHORN; TISSUE=EMBRYONIC BRAIN;			
RX	MEDLINE; 94340732.			
RA	SERAFINI T., KENNEDY T.E., GALKO M.J., MIRZAYAN C., JESSELL T.M.,			
RA	TESSIER-LAVIGNE M.;			
RL	CELL 78:409-424(1994).			
CC	-1- FUNCTION: NETRINS CONTROL GUIDANCE OF CNS COMMISSURAL AXONS AND PERIPHERAL MOTOR AXONS.			
CC	-1- SUBCELLULAR LOCATION: EXTRACELLULAR.			
CC	-1- SIMILARITY: CONTAINS ONE LAMININ N-TERMINAL DOMAIN (DOMAIN VI).			
CC	-1- SIMILARITY: CONTAINS 3 LAMININ EGF-LIKE DOMAINS.			
CC	-1- SIMILARITY: CONTAINS ONE C345C DOMAIN.			
CC	-1- SIMILARITY: TO CAENORHABDITIS ELEGANS UNC-6 PROTEIN.			
DR	EWBL; L34549; G529419;			
DR	PROSITE; PS00222; EGF_1; 2.			
DR	PROSITE; PS001248; LAMININ_TYPE_EGF; 3.			
KW	GLYCOPROTEIN; EXTRACELLULAR MATRIX; SIGNAL; LAMININ EGF-LIKE DOMAIN;			
KW	REPEAT.			
FT	SIGNAL	1	25	POTENTIAL.
FT	CHAIN	26	606	NETRIN-1.
FT	DOMAIN	26	286	LAMININ N-TERMINAL (DOMAIN VI).
FT	DOMAIN	287	435	3 X LAMININ EGF-LIKE REPEATS (LAMININ DOMAIN V-LIKE).
FT	DOMAIN	287	342	LAMININ EGF-LIKE 1.
FT	DOMAIN	343	405	LAMININ EGF-LIKE 2.
FT	DOMAIN	406	455	LAMININ EGF-LIKE 3.
FT	DOMAIN	456	606	C345C (DOMAIN C).
FT	SITE	532	534	CELL ATTACHMENT SITE (POTENTIAL).
FT	DISULFID	121	154	BY SIMILARITY.
FT	DISULFID	287	296	BY SIMILARITY.
FT	DISULFID	289	306	BY SIMILARITY.
FT	DISULFID	308	317	BY SIMILARITY.
FT	DISULFID	320	340	BY SIMILARITY.
FT	DISULFID	343	352	BY SIMILARITY.
FT	DISULFID	345	370	BY SIMILARITY.
FT	DISULFID	373	382	BY SIMILARITY.
FT	DISULFID	385	403	BY SIMILARITY.
FT	DISULFID	406	418	BY SIMILARITY.
FT	DISULFID	408	425	BY SIMILARITY.
FT	DISULFID	427	436	BY SIMILARITY.
FT	DISULFID	439	453	BY SIMILARITY.


```

[2]
RN      SEQUENCE OF 177-500 FROM N.A.
RX      MEDLINE: 89050091.
RA      ISHIDA N., TAWARAGI Y., INUZUKA C., SUGITA O., KUBOTA I.,
RA      NAKAZATO H., NOGUCHI T., SASSA S.;
RL      BIOCHEM. BIOPHYS. RES. COMMUN. 156:681-688(1988).
CC      -!- FUNCTION: CYTOCHROMES P450 ARE A GROUP OF HEME-THIOLATE
CC      MONOOXYGENASES. IN LIVER MICROSOMES, THIS ENZYME IS INVOLVED IN AN
CC      NAOPH-DEPENDENT ELECTRON TRANSPORT PATHWAY. IT OXIDIZES A VARIETY
CC      OF STRUCTURALLY UNRELATED COMPOUNDS, INCLUDING STEROIDS, FATTY
CC      ACIDS, AND XENOBIOTICS.
CC      -!- CATALYTIC ACTIVITY: RH + REDUCED FLAVOPROTEIN + O(2) -> ROH +
CC      OXIDIZED FLAVOPROTEIN + H(2)O.
CC      -!- SUBCELLULAR LOCATION: MEMBRANE-BOUND.
CC      -!- INDUCTION: P450 CAN BE INDUCED TO HIGH LEVELS IN LIVER AND OTHER
CC      TISSUES BY VARIOUS FOREIGN COMPOUNDS, INCLUDING DRUGS, PESTICIDES,
CC      AND CARCINOGENS.
CC      -!- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.
DR      EMBL: X52029; G57816; -.
DR      EMBL: M22331; G203830; -.
DR      PIR: D31579; D31579.
DR      PIR: S16873; S16873.
DR      PROSITE: PS00086; CYTOCHROME_P450; 1.
KW      OXIDOREDUCTASE; MONOOXYGENASE; ELECTRON TRANSPORT; MEMBRANE; HEME;
KW      MICROSOME.
FT      BINDING          446      446      HEME (BY SIMILARITY).
SQ      SEQUENCE      500 AA; 56697 MW; 429F82E7 CRC32;

Query Match          4.7%; Score 105; DB 1; Length 500;
Best Local Similarity 24.7%; Pred. No. 6,13e-03;
Matches 18; Conservative 28; Mismatches 21; Indels 6; Gaps 3

Db      160 ARCLCAAFADHSGFFPS-PTLLDKAVCNVCIASLLFACR-FEYNDPRFIRLL--DLKQDT 215
      :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
QY      15 SHC-CLG-SARGLEFLFGQDFSYKRSNCKPIPNQLCHGIEYQNMLPLNLGHETMKV 72
      :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::

Db      216 LEEESGFLPMLN 228
      :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
QY      73 LEQAGAWIPLVMK 85
      :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::

RESULT      6
ID      CPDI_RAT
AC      Q64680;
DT      01-NOV-1997 (REL. 35, CREATED)
DT      01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)
DE      01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE      CYTOCHROME P450 IID18 (EC 1.14.14.1) (P450 2D-29/2D-35).
GN      CYP2D18 OR CYP2D-18.
OS      RATTUS NORVEGICUS (RAT).
OS      EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC      EUTHERIA; RODENTIA.
[1]
RN      SEQUENCE FROM N.A.
RX      MEDLINE: 95251650.
RA      KAWASHIMA H., STROBEL H.W.;
RL      BIOCHEM. BIOPHYS. RES. COMMUN. 209:535-540(1995).
CC      -!- FUNCTION: CYTOCHROMES P450 ARE A GROUP OF HEME-THIOLATE
CC      MONOOXYGENASES. IN LIVER MICROSOMES, THIS ENZYME IS INVOLVED IN AN
CC      NAOPH-DEPENDENT ELECTRON TRANSPORT PATHWAY. IT OXIDIZES A VARIETY
CC      OF STRUCTURALLY UNRELATED COMPOUNDS, INCLUDING STEROIDS, FATTY
CC      ACIDS, AND XENOBIOTICS.
CC      -!- CATALYTIC ACTIVITY: RH + REDUCED FLAVOPROTEIN + O(2) -> ROH +
CC      OXIDIZED FLAVOPROTEIN + H(2)O.
CC      -!- SUBCELLULAR LOCATION: MEMBRANE-BOUND.
CC      -!- TISSUE SPECIFICITY: BRAIN.
CC      -!- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.
DR      EMBL: U48220; G1200518; -.
DR      EMBL: U48219; G1200516; -.
DR      PROSITE: PS00086; CYTOCHROME_P450; 1.
KW      OXIDOREDUCTASE; MONOOXYGENASE; ELECTRON TRANSPORT; MEMBRANE; HEME;
KW      MICROSOME.
FT      BINDING          446      446      HEME (BY SIMILARITY).

```

```

OS MUS MUSCULUS (MOUSE) .
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC EUTHERIA; RODENTIA.
[1]
RN SEQUENCE FROM N.A.
RP MEDLINE; 89384612.
RX ALBERTA D.A.; CREWS C.M., SWEET L.J., BANKSTON L., JONES S.W.,
RA ERIKSON R.L.;
RL MOL. CELL. BIOL. 9:3850-3859(1989).
CC -!- FUNCTION: PHOSPHORYLATES A WIDE RANGE OF SUBSTRATES INCLUDING
CC RIBOSOMAL PROTEIN S6.
CC -!- TISSUE SPECIFICITY: INTESTINE, THYMUS, AND LUNG.
CC -!- SIMILARITY: THE N-TERMINUS SHOWS HIGH SEQUENCE SIMILARITY TO THE
CC REGIONS OF PROTEIN KINASE C, THE CATALYTIC SUBUNIT OF CAMP-
CC DEPENDENT PROTEIN KINASE, AND CAMP-DEPENDENT PROTEIN KINASE. THE
CC REMINDER OF S6 KINASE II SHOWS HIGH SEQUENCE SIMILARITY TO THE
CC CATALYTIC SUBUNIT OF PHOSPHORYLASE B KINASE.
CC EMBL; M28489; G556322; -.
DR DR PIR; B32571; B32571.
DR HSSP; P05132; ICTP.
DR MGD; MGI:104558; RPS6KAL.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 2.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 2.
DR PROSITE; PS00111; PROTEIN_KINASE_DOM; 2.
DR TRANSFERASE; SERINE/THREONINE-PROTEIN KINASE; ATP-BINDING;
KW REPEAT; MULTIGENE FAMILY.
FT FT DOMAIN 52 310 PROTEIN KINASE 1.
FT FT DOMAIN 407 664 PROTEIN KINASE 2.
FT FT NP_BIND 68 76 ATP (BY SIMILARITY).
FT FT BINDING 94 94 ATP (BY SIMILARITY).
FT FT ACT_SITE 187 187 BY SIMILARITY.
FT FT NP_BIND 413 421 ATP (BY SIMILARITY).
FT FT BINDING 436 436 ATP (BY SIMILARITY).
FT FT ACT_SITE 524 524 BY SIMILARITY.
FT SEQ SEQUENCE 724 AA; 81594 MW; F961A261 CRC32;

Query Match 4.4%; Score 99; DB 1; Length 724;
Best Local Similarity 35.2%; Pred. No. 5.67e-02;
Matches 19; Conservative 13; Mismatches 18; Indels 4; Gaps

Db 489 KI-LRQKFFSREASFLVLTISKIVTVLHSGQVYVHRDLKPSNLYVDESGNPEC 541
|| :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::
QY 197 KIKVKEITYINRDKTIILETKSKTI-Y-KLVGVSRDLKKS-VLVLKDSLQCTC 247

RESULT 9
ID VG07_BPT4 STANDARD; PRT; 1032 AA.
AC P19061.
DT 01-NOV-1990 (REL. 16, CREATED)
DT 01-NOV-1990 (REL. 16, LAST SEQUENCE UPDATE)
DE 01-NOV-1990 (REL. 16, LAST ANNOTATION UPDATE)
DE BASEPLATE STRUCTURAL PROTEIN GP7.
GN 7.
OC BACTERIOPHAGE T4.
OC VIRIDAE; DS-DNA NONENVELOPED VIRUSES; MYOVIRIDAE.
[1]
RN SEQUENCE FROM N.A.
RP STRAIN-D;
RX MEDLINE; 90384864.
RL EFIMOV V.P., PRILIPOV A.G., MESVANYZHINOV V.V.;
RL NUCLEIC ACIDS RES 18:5313-5313(1990).
CC -!- FUNCTION: STRUCTURAL COMPONENT OF THE BASEPLATE.
DR EMBL; X15907; G15323; -.
DR PIR; JQ0657; G7BPT4.
DR STRUCTURAL PROTEIN.
KW SEQUENCE 1032 AA; 119214 MW; 39F5B2D1 CRC32;

Query Match 4.4%; Score 98; DB 1; Length 1032;
Best Local Similarity 39.0%; Pred. No. 8.13e-02;
Matches 16; Conservative 8; Mismatches 15; Indels 2; Gaps

Db 368 MDSICDKVFALIIIEVETLANPRTSKII-DSADKGIYVLN 407
||||: ||| :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::

```

CC SUBSEQUENT ELIMINATION OF EXOGENOUS COMPOUNDS.

OS ACINETOBACTER SP. (STRAIN ATCC 17924 / CIP 70.12).

CC PROKARYOTA; GRACILICUTES; SCOTOBACTERIA; AEROBIC RODS AND COCCI;
 CC NEISSERIACEAE.
 CC [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE: 9708999.
 RA YAMAMOTO S., HARAYAMA S.;
 RL INT. J. SYST. BACTERIOL. 46:506-511(1996).
 CC -!- FUNCTION: DNA GYRASE NEGATIVELY SUPERCOILS CLOSED CIRCULAR DOUBLE-
 CC STRANDED DNA IN AN ATP-DEPENDANT MANNER AND ALSO CATALYZES THE
 CC INTERCONVERSION OF OTHER TOPOLOGICAL ISOMERS OF DOUBLE-STRANDED
 CC DNA RINGS, INCLUDING CATENANES AND KNOTTED RINGS.
 CC -!- CATALYTIC ACTIVITY: ATP-DEPENDENT BREAKAGE, PASSAGE AND REJOINING
 CC OF DOUBLE-STRANDED DNA.
 CC -!- SUBUNIT: MADE UP OF TWO CHAINS. THE A CHAIN IS RESPONSIBLE FOR DNA
 CC BREAKAGE AND REJOINING; THE B CHAIN CATALYZES ATP HYDROLYSIS. THE
 CC ENZYME FORMS AN A2B2 TETRAMER.
 DR EMBL: D73436; G1322131; -.
 DR EMBL: D73421; G1322101; -.
 DR PROSITE: PS00177; TOPOISOMERASE_II; 1.
 KW TOPOISOMERASE; ISOMERASE; ATP-BINDING; ANTIBIOTIC RESISTANCE.
 FT NON_TER 1
 FT NON_CONS 116 117
 FT NON_TER 216 216
 SQ SEQUENCE 216 AA; 23775 MW; E572EEE0 CRC32;
 Query Match 4.0%; Score 90; DB 1; Length 216;
 Best Local Similarity 26.3%; Pred. No. 1.31e+00;
 Matches 20; Conservative 23; Mismatches 26; Indels 7; Gaps 7;
 Db 79 LARRRLELFLNAGVRVIRDRERVALEHIFDVLGLSEKSL-DIAGLPGLKA-DCQEKD 136
 Qy 196 LKIKVKEITYINRDKILL-ETK-S-KTIYKLN-GVSEKDLKSVLWKLKDSLOCTCEMN 251
 Db 137 PALSELXIVGDSAGG 152
 Qy 252 -DINAPYLVMQKOGG 266
 RESULT 14
 ID VIRF_SHIDY STANDARD; PRT; 262 AA.
 AC Q04248;
 DT 01-JUN-1994 (REL. 29, CREATED)
 DT 01-JUN-1994 (REL. 29, LAST SEQUENCE UPDATE)
 DT 01-NOV-1995 (REL. 32, LAST ANNOTATION UPDATE)
 DE VIRULENCE REGULON TRANSCRIPTIONAL ACTIVATOR VIRF.
 GN VIRF.
 OS SHIGELLA DYSENTERIAE, SHIGELLA FLEXNERI, AND SHIGELLA SONNEI.
 OG PLASMID 210 KB INVASION.
 CC PROKARYOTA; GRACILICUTES; SCOTOBACTERIA; FACULTATIVELY ANAEROBIC RODS;
 CC ENTEROBACTERIACEAE.
 CC [1]
 RP SEQUENCE FROM N.A.
 RC SPECIES=S.DYSENTERIAE; STRAIN=CG097;
 RA YAO R., REDDY L.V., PALCHANDHURI S.;
 RL SUBMITTED (MAR-1991) TO EMBL/GENBANK/DBJ DATA BANKS.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC SPECIES=S.FLEXNERI; STRAIN=2A;
 RX MEDLINE: 87032409.
 RA SAKAI T., SASAKAWA C., MAKINO S., YOSHIKAWA M.;
 RL INFECT. IMMUN. 54:395-402(1986).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC SPECIES=S.SONNEI;
 RX MEDLINE: 89212891.
 RA KATO J.I., ITO K.I., NAKAMURA A., WATANABE H.;
 RL INFECT. IMMUN. 57:1391-1398(1989).
 RN [4]
 RP SIMILARITY TO ARAC FAMILY.
 RX MEDLINE: 92326642.
 RA DORMAN C.J.;
 RL MOL. MICROBIOL. 6:1575-1575(1992).
 CC -!- FUNCTION: TRANSCRIPTIONAL ACTIVATOR OF THE VIRB GENE WHICH IS

CC ITSELF AN ACTIVATOR OF THE IPAABCD VIRULENCE REGULON.
 CC -!- SIMILARITY: BELONGS TO THE ARAC/XVLS FAMILY OF TRANSCRIPTIONAL
 CC REGULATORS.
 DR EMBL: X58464; G46954; -.
 DR EMBL: M29172; G152801; -.
 DR EMBL: X16661; G47067; -.
 DR PIR: S14646; S14646.
 DR PIR: A47605; A47605.
 DR PIR: A60105; A60105.
 DR PROSITE: PS00041; HTH_ARAC_FAMILY_1; 1.
 DR PROSITE: PS01124; HTH_ARAC_FAMILY_2; 1.
 KW PLASMID; VIRULENCE; TRANSCRIPTION REGULATION; ACTIVATOR; DNA-BINDING.
 FT DNA_BIND 177 196 H-T-H MOTIF (BY SIMILARITY).
 SQ SEQUENCE 262 AA; 30549 MW; 11339B38 CRC32;
 Query Match 4.0%; Score 89; DB 1; Length 262;
 Best Local Similarity 25.5%; Pred. No. 1.83e+00;
 Matches 13; Conservative 21; Mismatches 15; Indels 2; Gaps 2;
 Db 37 TLAIDEGQAFIERNTQINVSIKSDSINPFETISIDRNLLLSIIRMEPI 87
 Qy 195 ALKIKVKEITYINRDKILL-ETK-S-KTIYKLNQVS-ERDLKKSVLWKDSL 243
 RESULT 15
 ID KSGA_MYCPN STANDARD; PRT; 263 AA.
 AC P75113;
 DT 01-NOV-1997 (REL. 35, CREATED)
 DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)
 DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
 DE DIMETHYLADENOSINE TRANSFERASE (EC 2.1.1.-) (S-ADENOSYLMETHIONINE-6-N',
 DE N'-ADENOSYL(RRNA) DIMETHYLTRANSFERASE) (16S RNA DIMETHYLASE) (HIGH
 DE LEVEL KASUGAMYCIN RESISTANCE PROTEIN KSGA) (KASUGAMYCIN
 DE DIMETHYLTRANSFERASE).
 GN KSGA.
 OS MYCOPLASMA PNEUMONIAE.
 CC PROKARYOTA; TENERICUTES; MOLLICUTES; MYCOPLASMA; MYCOPLASMATALES;
 CC MYCOPLASMATAEAE.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 29342 / M129;
 RX MEDLINE: 97105885.
 RA HIMMELREICH R., HILBERT H., PLAGENS H., PIRKL E., LI B.-C.,
 RA HERRMANN R.;
 RL NUCLEIC ACIDS RES. 24:4420-4449(1996).
 CC -!- FUNCTION: SPECIFICALLY DIMETHYLATES TWO ADJACENT ADENOSINES IN THE
 CC LOOP OF A CONSERVED HAIRPIN NEAR THE 3'-END OF 16S RNA IN THE 30S
 CC PARTICLE. ITS INACTIVATION LEADS TO KASUGAMYCIN RESISTANCE
 CC (BY SIMILARITY).
 CC -!- SIMILARITY: STRONG, TO OTHER BACTERIAL KSGA.
 CC -!- SIMILARITY: BELONGS TO THE RNA ADENINE N-6-METHYLTRANSFERASE
 CC FAMILY.
 DR EMBL: AE000017; G1673824; -.
 DR PROSITE: PS01131; RNA_A_DIMETH; 1.
 KW RNA PROCESSING; TRANSFERASE; METHYLTRANSFERASE;
 KW ANTIBIOTIC RESISTANCE.
 SQ SEQUENCE 263 AA; 29794 MW; 357AA00D CRC32;
 Query Match 4.0%; Score 90; DB 1; Length 263;
 Best Local Similarity 37.0%; Pred. No. 1.31e+00;
 Matches 20; Conservative 11; Mismatches 20; Indels 3; Gaps 3;
 Db 52 KLQLPYHGIEL-DKRLAEYLLVNEILTEQLTGIDGALKONLQYFDPDTPLLCG 104
 Qy 46 NLQL-CHGIEYQNMRLPN-LLGHETMKVELEGAGAWIPLVKQCHPDTKKFLCS 97
 Search completed: Tue Aug 4 10:34:10 1998
 Job time : 26 secs.


```

RESULT 2
ID O35297 PRELIMINARY; PRT: 295 AA.
AC O35297;
DT 01-JAN-1998 (TREMBLREL. 05, CREATED)
DT 01-JAN-1998 (TREMBLREL. 05, LAST SEQUENCE UPDATE)
DT 01-JAN-1998 (TREMBLREL. 05, LAST SEQUENCE UPDATE)
DE SECRETED APOPTOSIS RELATED PROTEIN 1.
GN SARPI.
OS MUS MUSCULUS (MOUSE).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC EUTHERIA; RODENTIA.
RN [1]
RP SEQUENCE FROM N.A.
RA MELKONYAN H., CHANG W.C., SHAPIRO J.P., MAHADEVAPPA M., FITZPATRICK P.A.,
RA FITZPATRICK P.A., KIEFER M.C., TOMEI D.L., UMANSKY S.R.;
RL SUBMITTED (SEP-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL; AF017989; G2415421; -.
SQ SEQUENCE 295 AA; 33483 MW; 0E4C8A8E CRC32;

Query Match 97.9%; Score 2188; DB 10; Length 295;
Best Local Similarity 97.6%; Pred. No. 0.00e+00;
Matches 288; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

Db 1 MPRGPASLLLVASHCCLSGARGFLFGQPDFSYKRNCKRPIPANLQCHGIEYQNRL 60
QY 1 MLOQPGSLLLVFLASHCCLSGARGFLFGQPDFSYKRNCKRPIPANLQCHGIEYQNRL 60
Db 61 PNLLGHETMKEVLEQAGAWIPLVMKQCHPDTKKFLCSLFAFVCLDDLDDETQPCHSVCQ 120
QY 61 PNLLGHETMKEVLEQAGAWIPLVMKQCHPDTKKFLCSLFAFVCLDDLDDETQPCHSVCQ 120
Db 121 VKDRCAPVMSAFGFPWPDMLCEDRFPODNDLCIPLASSDHLPLATEAPKVCACKTKNE 180
QY 121 VKDRCAPVMSAFGFPWPDMLCEDRFPODNDLCIPLASSDHLPLATEAPKVCACKTKNE 180
Db 181 DDNDIMETLCKNDPALKIKVKEIYINRDKIILETSKTIYKLVGVSERDLKKSVMWLK 240
QY 181 DDNDIMETLCKNDPALKIKVKEIYINRDKIILETSKTIYKLVGVSERDLKKSVMWLK 240
Db 241 DSLOQCTCEEMNDINAPYLVMGQKGGELVITSVKRWQKQREFKRISIRKLQ 295
QY 241 DSLOQCTCEEMNDINAPYLVMGQKGGELVITSVKRWQKQREFKRISIRKLQ 295

RESULT 3
ID P97299 PRELIMINARY; PRT: 295 AA.
AC P97299;
DT 01-MAY-1997 (TREMBLREL. 03, CREATED)
DT 01-MAY-1997 (TREMBLREL. 03, LAST SEQUENCE UPDATE)
DT 01-JAN-1998 (TREMBLREL. 05, LAST SEQUENCE UPDATE)
DE STROMAL CELL DERIVED FACTOR 5 (SDF5).
GN SDF5.
OS MUS MUSCULUS (MOUSE).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC EUTHERIA; RODENTIA.
RN [1]
RP SEQUENCE FROM N.A.
RA SHIROZU M., TADA H., TASHIRO K., NAKAMURA T., LOPEZ N.D.,
RA NAZARENA M., HAMADA T., SATO T., NAKANO T., HONJO T.;
RL GENOMICS 37:273-280(1996).
DR EMBL; D50462; G1747302; -.
DR MGD; MGI:108078; SDF5.
SQ SEQUENCE 295 AA; 33501 MW; 329888C8 CRC32;

Query Match 97.8%; Score 2185; DB 10; Length 295;
Best Local Similarity 97.6%; Pred. No. 0.00e+00;
Matches 288; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

Db 1 MPRGPASLLLVASHCCLSGARGFLFGQPDFSYKRNCKRPIPANLQCHGIEYQNRL 60
QY 1 MLOQPGSLLLVFLASHCCLSGARGFLFGQPDFSYKRNCKRPIPANLQCHGIEYQNRL 60
Db 61 PNLLGHETMKEVLEQAGAWIPLVMKQCHPDTKKFLCSLFAFVCLDDLDDETQPCHSVCQ 120
QY 61 PNLLGHETMKEVLEQAGAWIPLVMKQCHPDTKKFLCSLFAFVCLDDLDDETQPCHSVCQ 120

```

```

QY 61 PNLLGHETMKEVLEQAGAWIPLVMKQCHPDTKKFLCSLFAFVCLDDLDDETQPCHSVCQ 120
Db 121 VKDRCAPVMSAFGFPWPDMLCEDRFPODNDLCIPLASSDHLPLATEAPKVCACKTKNE 180
QY 121 VKDRCAPVMSAFGFPWPDMLCEDRFPODNDLCIPLASSDHLPLATEAPKVCACKTKNE 180
Db 181 DDNDIMETLCKNDPALKIKVKEIYINRDKIILETSKTIYKLVGVSERDLKKSVMWLK 240
QY 181 DDNDIMETLCKNDPALKIKVKEIYINRDKIILETSKTIYKLVGVSERDLKKSVMWLK 240
Db 241 DSLOQCTCEEMNDINAPYLVMGQKGGELVITSVKRWQKQREFKRISIRKLQ 295
QY 241 DSLOQCTCEEMNDINAPYLVMGQKGGELVITSVKRWQKQREFKRISIRKLQ 295

RESULT 4
ID O14778 PRELIMINARY; PRT: 206 AA.
AC O14778;
DT 01-JAN-1998 (TREMBLREL. 05, CREATED)
DT 01-JAN-1998 (TREMBLREL. 05, LAST SEQUENCE UPDATE)
DT 01-JAN-1998 (TREMBLREL. 05, LAST ANNOTATION UPDATE)
DE SECRETED APOPTOSIS RELATED PROTEIN 1 (FRAGMENT).
GN SARPI.
OS HOMO SAPIENS (HUMAN).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC EUTHERIA; PRIMATES.
RN [1]
RP SEQUENCE FROM N.A.
RA MELKONYAN H., CHANG W.C., SHAPIRO J.P., MAHADEVAPPA M., FITZPATRICK P.A.,
RA KIEFER M.C., TOMEI D.L., UMANSKY S.R.;
RL SUBMITTED (SEP-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL; AF017986; G2415415; -.
FT NON_TER 206
SQ SEQUENCE 206 AA; 23215 MW; F465B68B CRC32;

Query Match 68.8%; Score 1538; DB 2; Length 206;
Best Local Similarity 98.6%; Pred. No. 0.00e+00;
Matches 205; Conservative 0; Mismatches 1; Indels 2; Gaps 1;

Db 1 MLOQPGSLLLVFLASHCCLSGARGFLFGQPDFSYKRNCKRPIPANLQCHGIEYQNRL 60
QY 1 MLOQPGSLLLVFLASHCCLSGARGFLFGQPDFSYKRNCKRPIPANLQCHGIEYQNRL 60
Db 61 PNLLGHETMKEVLEQAGAWIPLVMKQCHPDTKKFLCSLFAFVCLDDLDDETQPCHSVCQ 120
QY 61 PNLLGHETMKEVLEQAGAWIPLVMKQCHPDTKKFLCSLFAFVCLDDLDDETQPCHSVCQ 120
Db 121 VKDRCAPVMSAF--PWPDMLCEDRFPODNDLCIPLASSDHLPLATEAPKVCACKTKNE 178
QY 121 VKDRCAPVMSAFGFPWPDMLCEDRFPODNDLCIPLASSDHLPLATEAPKVCACKTKNE 180
Db 179 DDNDIMETLCKNDPALKIKVKEIYINR 206
QY 181 DDNDIMETLCKNDPALKIKVKEIYINR 208

RESULT 5
ID O14779 PRELIMINARY; PRT: 314 AA.
AC O14779;
DT 01-JAN-1998 (TREMBLREL. 05, CREATED)
DT 01-JAN-1998 (TREMBLREL. 05, LAST SEQUENCE UPDATE)
DT 01-JAN-1998 (TREMBLREL. 05, LAST ANNOTATION UPDATE)
DE SECRETED APOPTOSIS RELATED PROTEIN 2.
GN SARPI.
OS HOMO SAPIENS (HUMAN).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC EUTHERIA; PRIMATES.
RN [1]
RP SEQUENCE FROM N.A.
RA MELKONYAN H., PROCHAZKA V., CHANG W.C., UMANSKY S.R.;
RL SUBMITTED (SEP-1997) TO EMBL/GENBANK/DBJ DATA BANKS.

```

```
DR EMBL; AF017987; G2415417; -.
SQ SEQUENCE 314 AA; 35411 MW; 61E261B5 CRC32;

Query Match
Best Local Similarity 40.9%; Score 913; DB 2; Length 314;
Matches 112; Conservative 67; Mismatches 77; Indels 7; Gaps 5;

Db 51 FTKPPQCVDPADLRCHNVGKRMVLPNLLHETMAEVKQOASSWVPLLNKNCHAGTQ 110
QY 33 FSKRSNCKPIPANLQCHGIEYQNMRLPNLLGHETMKVELEQAGAWIPLVMKQCHPDTK 92

Db 111 VFLCSLFAPVCLD---RPIYPCRWLCEAVRDSCEPVMQFFGYWPMLKCDKFP-EGDVC 166
QY 93 KFLCSLFAPVCLDLDDETTOPCHSLCVQVQKRCAPVMSAFGFPWPDMLCEDRFPQDNDLC 152

Db 167 IAMTPNPTEASKPGQTTVPCPCDNLKSEA-IIIEHLCASEFALRMKIKVEKKENDKKI 225
QY 153 IPLASSDHLHPATEAPKVCCEACKNKNDNDNDIMETLCKNDFAIKKVEITYINRDTKI 212

Db 226 V-PKKKKPL-KLGPPIKKDLKLVLYLKGADCPCHOLDNLSHHFLIMGRKVKSOYLLTA 283
QY 213 ILETSKTIYKLVGVSERDLKKSVLWKLDSLOCTCEEMNDINAPYLVMGOKGOGELVITS 272

Db 284 IHKWDKKNKEFKNMKKNHEC 306
QY 273 VKRWQKGREFRISRIRKLOC 295

RESULT 6
ID 000546; PRELIMINARY; PRT; 313 AA.
AC 000546;
DT 01-JUL-1997 (TREMBLREL. 04, CREATED)
DT 01-JUL-1997 (TREMBLREL. 04, LAST SEQUENCE UPDATE)
DT 01-JUL-1997 (TREMBLREL. 04, LAST ANNOTATION UPDATE)
DE SECRETED FRIZZLED-RELATED PROTEIN.
OS HOMO SAPIENS (HUMAN).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC EUTHERIA; PRIMATES.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=LUNG;
RA FINCH P.W., HE X., KELLEY M.J., UREN A., SCHAUDIES R.P.,
RA POPESCU N.C., RUDIKOFF S., AARONSON S.A., VARMUS H.E., RUBIN J.S.;
RL SUBMITTED (JUN-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL; AF001900; G2213819; -.
SQ SEQUENCE 313 AA; 35314 MW; A3CD04FF CRC32;

Query Match
Best Local Similarity 40.8%; Score 911; DB 2; Length 313;
Matches 112; Conservative 67; Mismatches 77; Indels 7; Gaps 5;

Db 50 FTKPPQCVDPADLRCHNVGKRMVLPNLLHETMAEVKQOASSWVPLLNKNCHAGTQ 109
QY 33 FSKRSNCKPIPANLQCHGIEYQNMRLPNLLGHETMKVELEQAGAWIPLVMKQCHPDTK 92

Db 110 VFLCSLFAPVCLD---RPIYPCRWLCEAVRDSCEPVMQFFGYWPMLKCDKFP-EGDVC 165
QY 93 KFLCSLFAPVCLDLDDETTOPCHSLCVQVQKRCAPVMSAFGFPWPDMLCEDRFPQDNDLC 152

Db 166 IAMTPNPTEASKPGQTTVPCPCDNLKSEA-IIIEHLCASEFALRMKIKVEKKENDKKI 224
QY 153 IPLASSDHLHPATEAPKVCCEACKNKNDNDNDIMETLCKNDFAIKKVEITYINRDTKI 212

Db 225 V-PKKKKPL-KLGPPIKKDLKLVLYLKGADCPCHOLDNLSHHFLIMGRKVKSOYLLTA 282
QY 213 ILETSKTIYKLVGVSERDLKKSVLWKLDSLOCTCEEMNDINAPYLVMGOKGOGELVITS 272

Db 283 IHKWDKKNKEFKNMKKNHEC 305
QY 273 VKRWQKGREFRISRIRKLOC 295

RESULT 7
```

```
ID 019116; PRELIMINARY; PRT; 308 AA.
AC 019116;
DT 01-JAN-1998 (TREMBLREL. 05, CREATED)
DT 01-JAN-1998 (TREMBLREL. 05, LAST SEQUENCE UPDATE)
DT 01-JAN-1998 (TREMBLREL. 05, LAST ANNOTATION UPDATE)
DE FRAZ PRECURSOR.
OS BOS TAURUS (BOVINE).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC EUTHERIA; ARTIODACTYLA.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=HEART;
RA DUPLAA C., D'AMORE P.A.;
RL SUBMITTED (AUG-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL; U85945; G2337937; -.
KW SIGNAL.
FT SIGNAL.
SQ SEQUENCE 308 AA; 34763 MW; 9B74A92F CRC32;

Query Match
Best Local Similarity 40.4%; Score 902; DB 4; Length 308;
Matches 111; Conservative 66; Mismatches 79; Indels 7; Gaps 5;

Db 45 FTKPPQCVDPADLRCHNVGKRMVLPNLLHETMAEVKQOASSWVPLLNKNCHIGTQ 104
QY 33 FSKRSNCKPIPANLQCHGIEYQNMRLPNLLGHETMKVELEQAGAWIPLVMKQCHPDTK 92

Db 105 VFLCSLFAPVCLD---RPIYPCRWLCEAVRDSCEPVMQFFGYWPMLKCDKFP-EGDVC 160
QY 93 KFLCSLFAPVCLDLDDETTOPCHSLCVQVQKRCAPVMSAFGFPWPDMLCEDRFPQDNDLC 152

Db 161 IAMTPNPTEASKPGQTTVPCPCDNLKSEA-IIIEHLCASEFALRMKIKVEKKENDKKI 219
QY 153 IPLASSDHLHPATEAPKVCCEACKNKNDNDNDIMETLCKNDFAIKKVEITYINRDTKI 212

Db 220 V-PKKKKPL-KLGPPIKKDLKLVLYLKGADCPCHOLDNLSHHFLIMGRKVKSOYLLTA 277
QY 213 ILETSKTIYKLVGVSERDLKKSVLWKLDSLOCTCEEMNDINAPYLVMGOKGOGELVITS 272

Db 278 IHKWDKKNKEFKNMKKNHEC 300
QY 273 VKRWQKGREFRISRIRKLOC 295

RESULT 8
ID 014780; PRELIMINARY; PRT; 317 AA.
AC 014780;
DT 01-JAN-1998 (TREMBLREL. 05, CREATED)
DT 01-JAN-1998 (TREMBLREL. 05, LAST SEQUENCE UPDATE)
DT 01-JAN-1998 (TREMBLREL. 05, LAST ANNOTATION UPDATE)
DE SECRETED APOPTOSIS RELATED PROTEIN 3.
GN SARP3.
OS HOMO SAPIENS (HUMAN).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC EUTHERIA; PRIMATES.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=PANCREAS;
RA MELKONTAN H., CHANG W.C., SHAPIRO J.P., MAHADEVAPPA M., FITZPATRICK P.A.,
RA KIEFER M.C., TOMBI D.L., UNANSKY S.R.;
RL SUBMITTED (SEP-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL; AF017988; G2415419; -.
SQ SEQUENCE 317 AA; 35577 MW; 2CD69373 CRC32;

Query Match
Best Local Similarity 40.3%; Score 901; DB 2; Length 317;
Matches 114; Conservative 58; Mismatches 74; Indels 6; Gaps 4;

Db 46 SYSKPPQCVDPADLRCHNVGKRMVLPNLLHETMAEVKQOASSWVPLLNKNCHSDTQ 105
QY 34 SY-KRSNCKPIPANLQCHGIEYQNMRLPNLLGHETMKVELEQAGAWIPLVMKQCHPDTK 92

Db 106 VFLCSLFAPVCLD---RPIYPCRWLCEAVRDSCEPVMQFFGYWPMLKCDKFP-EGDVC 162
```

93 KFLCSLFAVCLDLDLDTIOPCHSLCVQKDRCAPVMSAFGFPWDMLECDREFPDNDLC 152
Db 163 IAVQFG-HLPATAPVTKIAQC-EMHSADGLMQMGSSDFVVMKRIKEIKIENGDRKL 220
Qy 153 IPLASSDHLPPATEAPKVCCEACKNDDNDIMETLCKNDPALKIKVKEITYINRDTKI 212
Db 221 IGAQKKKLLKPGPKRDKTRKLVLMKNGAGCCPCOLDLAGSLVNGRVDGOLLMA 280
Qy 213 ILETASKIYKLVNGVSRDLKSLVWLKDSLOCTCEEMNDINAPLYVMGQKGGLVITS 272
Db 281 VYRWKKNKEMK 292
Qy 273 VKRWQKGQREFK 284

RESULT 9
ID O08861 PRELIMINARY; PRT; 314 AA.

AC O08861;
DT 01-JUL-1997 (TREMBREL. 04, CREATED)
DT 01-JUL-1997 (TREMBREL. 04, LAST SEQUENCE UPDATE)
DT 01-JUL-1997 (TREMBREL. 04, LAST ANNOTATION UPDATE)
DE SECRETED FRIZZLED RELATED PROTEIN SRP-1.
GN SRP1.
OS MUS MUSCULUS (MOUSE).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
RN EUTHERIA; RODENTIA.
RP SEQUENCE FROM N.A.
RA RATTNER A., HSIEH J.C., SMALLWOOD P.M., GILBERT D.J.,
RA COPELAND N.G., JENKINS N.A., NATHANS J.,
RL PROC. NATL. ACAD. SCI. U.S.A. 94:2859-2863(1997).
DR EMBL: U88566; G1946341; -.
SQ SEQUENCE 314 AA; 35327 MW; 5AA5CA27 CRC32;

Query Match 39.8%; Score 890; DB 10; Length 314;
Best Local Similarity 41.4%; Pred. No. 5.34e-193;
Matches 109; Conservative 67; Mismatches 80; Indels 7; Gaps 5;

51 FVTKPQCVDIPVLRILCHNNGYKKWVLPNLEHETMAEVKQASSVPLNKNCHMTQ 110
Qy 33 FSKRSNCKPIANLQCHGIEYQNMRLPNLLGHETMKEVLEQAGAWIPLVNKKOCHPDTK 92
Db 111 VFLCSLFAVCLD---RPIYPCWLCERVDSCFVPMOFFGFWPEMLKCDKFP-EGDVC 166
Qy 93 KFLCSLFAVCLDLDLDTIOPCHSLCVQKDRCAPVMSAFGFPWDMLECDREFPDNDLC 152
Db 167 IAMPPTNTTEASKQGTTCVPPCONELKSEA-IEHICASEFALRMKIKEYKENGDKKI 225
Qy 153 IPLASSDHLPPATEAPKVCCEACKNDDNDIMETLCKNDPALKIKVKEITYINRDTKI 212
Db 226 V-PKKKKPL-KLGPITKKELKALVFLKNGADCPCHOLDNLNHLNLMGRKVKSOYLLTA 283
Qy 213 ILETASKIYKLVNGVSRDLKSLVWLKDSLOCTCEEMNDINAPLYVMGQKGGLVITS 272
Db 284 IHKWDKKNKEFKNMRKNHNEC 306
Qy 273 VKRWQKGQREFKRSIRKLOC 295

RESULT 10
ID O42397 PRELIMINARY; PRT; 307 AA.

AC O42397;
DT 01-JAN-1998 (TREMBREL. 05, CREATED)
DT 01-JAN-1998 (TREMBREL. 05, LAST SEQUENCE UPDATE)
DT 01-JAN-1998 (TREMBREL. 05, LAST ANNOTATION UPDATE)
DE CRESCENT.
OS GALLUS GALLUS (CHICKEN).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; AVES; NEOGNATHAE;
RN GALLIFORMES.
RP SEQUENCE FROM N.A.
RA PFEFFER P.L., IZPISVA-BELMONTE J.C., DE ROBERTIS E.M.;

RL SUBMITTED (JUL-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL: AF006508; G2226372; -.
SQ SEQUENCE 307 AA; 34683 MW; 320A2243 CRC32;

Query Match 32.6%; Score 729; DB 12; Length 307;
Best Local Similarity 35.8%; Pred. No. 3.39e-151;
Matches 96; Conservative 68; Mismatches 94; Indels 10; Gaps 8;

Db 36 SYLRSSSCTAIPRSMALCYDYGISEMRIPNLLHEHETMPEVIOQSSSWLPLAREHPDA 95
Qy 34 SY-KRSN-CKPIANLQCHGIEYQNMRLPNLLGHETMKEVLEQAGAWIPLVNKKOCHPD 91
Db 96 RIFLCSLFAVCLDLDLDTIOPCHSLCVQKDRCAPVMSAFGFPWDMLECDREFPDNDL 151
Qy 92 KFLCSLFAVCLDLDLDTIOPCHSLCVQKDRCAPVMSAFGFPWDMLECDREFPDNDL 151
Db 153 CIAAVTDES-SSRRMPRASCKDELEAEASTAREILDNLCANDFTVKIRILKNTTTT 211
Qy 152 CIPLASSDHLPPATEAPKVCCEACK-NKNDNDNDIMETLCKNDPALKIKVKEITYINRD 209
Db 212 SDPDLPSRVVLKPGELLRTETIPGRLOQWLDIDATCAHNIMRGTHAGVVISGEVRS 271
Qy 210 TKIILETKSTIYKLVNGVSRDLKSLVWLKDSLOCTCEEMNDINAP-ILVMGQKQGE 267
Db 272 VYVKNAYAMOKNKNRNLHQAARRWKHRC 299
Qy 268 LVITSYKRWQKGQREFKRSIRKLOC 295

RESULT 11
ID O16147 PRELIMINARY; PRT; 568 AA.

AC O16147;
DT 01-JAN-1998 (TREMBREL. 05, CREATED)
DT 01-JAN-1998 (TREMBREL. 05, LAST SEQUENCE UPDATE)
DT 01-JAN-1998 (TREMBREL. 05, LAST ANNOTATION UPDATE)
DE MOM-5.
GN MOM-5.
OS CAENORHABDITIS ELEGANS.
OC EUKARYOTA; METAZOA; ACOELOMATES; NEMATODA; SECERNENTEA; RHABDITIDA.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL;
RA ROCHLEAU C.E., DOWNS W.D., LIN R., WITTMANN C., BEI Y., CHA Y.-H.,
RA ALI M., PRIESS J.R., MELLO C.C.;
RL CELL 90-707-716(1997).
DR EMBL: AF013953; G2463674; -.
SQ SEQUENCE 568 AA; 62892 MW; 8578BB4F CRC32;

Query Match 17.9%; Score 400; DB 3; Length 568;
Best Local Similarity 37.4%; Pred. No. 4.73e-68;
Matches 58; Conservative 33; Mismatches 54; Indels 10; Gaps 8;

Db 5 ILILFLFG-CLISADORLSTSTISSMNGFSTTR-KCEHI-T-IPACKNLDYNOTVFPNLL 59
Qy 8 LLLFLASHOCLGSA-R-GLFLFGQDFFSYKRSNCKPIANLQCHGIEYQNMRLPNLL 64
Db 60 GHTTQSEAGPAIAQFNPLIKVCSERILFLCTVYAPVC-TVLEKPIQPCRELCSAKNG 118
Qy 65 GHETMKEVLEQAGAWIPLVNKKOCHPDTKFELCSLFAVCLDLDLDTIOPCHSLCVQKDR 124
Db 119 CESLMKKFGQWPDQDCKNFPV-TDLCVGNKNSSE 152
Qy 125 CAPVMSAFGFPWDMLECDREFPDNDLCIPLASSD 159

RESULT 12
ID G61090 PRELIMINARY; PRT; 572 AA.

AC G61090;
DT 01-NOV-1996 (TREMBREL. 01, CREATED)
DT 01-NOV-1996 (TREMBREL. 01, LAST SEQUENCE UPDATE)
DT 01-JAN-1998 (TREMBREL. 05, LAST ANNOTATION UPDATE)
DE FRIZZLED HOMOLOG 7 (TRANSMEMBRANE RECEPTOR).
GN FZD7 OR FRIZZLED 7.


```

OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC EUTHERIA; RODENTIA.
[1]
RN SEQUENCE FROM N.A.
RX MEDLINE; 96224032.
RA WANG Y., MACKE J.P., ABELLA B.S., ANDREASSON K., WORLEY P.,
RL GILBERT D.J., COPELAND N.G., JENKINS N.A., NATHANS J.;
RJ J. BIOL. CHEM. 271:4468-4476(1996).
RR EMBL; U43321; G1151260; -
RS MGD; MGI:108460; FZD8.
RW TRANSMEMBRANE.
RX SEQUENCE 585 AA; 73215 MW; 8766F247 CRC32;
Query Match 17.6%; Score 394; DB 10; Length 685;
Best Local Similarity 43.8%; Pred. No. 1.39e-66;
Matches 56; Conservative 23; Mismatches 43; Indels 6; Gaps 6;

Db 40 VPLCKGIGYNTYMPNOFNHDTODEAGLEVHQFW-PLVEIQCSPLDKFFLCSTMTPICLE 98
QY 47 LQCHLGLEYQNMRLPNLGHETMKEV-LEQAGAWIPLVMKQCHPDTKKLCSLFAPVCLD 105
Db 99 DYKPLPPCSVGERAKAGCAPLMROYGFAPDPMRCDRLPEQGNPTLCMDYNRTD-LT 157
QY 106 DLDETIOCHSLCVQKDCAPVMSAFGFPWDMLECDREF-QDN'D-LCIPLASSDHL 162
Db 158 TAAPSPR 165
QY 163 PATEAPK 170

RESULT 15
ID Q14332 PRELIMINARY; PRT; 565 AA.
AC Q14332;
DT 01-NOV-1996 (TREMBLREL. 01, CREATED)
DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DT 01-NOV-1996 (TREMBLREL. 01, LAST ANNOTATION UPDATE)
DE FRIZZLED GENE PRODUCT.
DE FRIZZLED.
OS HOMO SAPIENS (HUMAN).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC EUTHERIA; PRIMATES.
RN [1]
RX SEQUENCE FROM N.A.
RC TISSUE=OVARY;
RX ZHAO Z., LEE C.C., BALDINI A., CASKEY C.T.;
RL GENOMICS 27:370-373(1995).
RR EMBL; L37882; G736679; -
RS SEQUENCE 565 AA; 63554 MW; 93392DA2 CRC32;
Query Match 17.3%; Score 386; DB 2; Length 565;
Best Local Similarity 37.4%; Pred. No. 1.24e-64;
Matches 46; Conservative 29; Mismatches 45; Indels 3; Gaps 3;

Db 41 PISIPCTDIAYNOTIMPNLGHTNOBDAGLEVHQFYPLVKVOCSPELRFLCSMTAPVC 100
QY 44 PANLQCHGIEYQNMRLPNLGHETMKEVLEQAGAWIPLVMKQCHPDTKKFLCSLFAVC 103
Db 101 -TVLEQAIIPCRSICERAROGCEALNNKFGQWPERLRCHEHPRHGAEOICVGNHSEDG 159
QY 104 LDDLEITQIPLCHSLCVQVKRCAPVMSAFGFPWDMLECDREF-QDND-LCIPLASSDHL 161
Db 160 APA 162
QY 162 LPA 164

Search completed: Tue Aug 4 10:35:55 1998
Job time : 87 secs.

```

Search completed: Tue Aug 4 10:35:55 1998
Job time : 87 secs.

THIS PAGE BLANK (USPTO)

QY 24 PANLQCHGIEYQNMRLPNLLGHETMKVEVLEQAGAWIPLVMKQCHPDTKKFLCSLFAFVC 83
Db 111 -tvldqaiapprscetargcealmnkkgfqpwrlercenfphgagelcvvgntsd 167
QY 84 LDDLETIQCHSLCVQVKDRCAPVMSAFGFPWDMLECDRFP-ODN-DLICIPLASSD 139

RESULT 2
ID W31274 standard; Protein: 685 AA.
AC W31274:
DT 27-APR-1998 (first entry)
DE Mouse frizzled-8 protein Mfz8 (Wnt receptor).
KW Wnt receptor; mouse frizzled-8 protein; Mfz8 gene;
KW signal transduction; cancer; cell growth; cell proliferation.
OS Mus musculus.
PN WO9739357-A1.
PD 23-OCT-1997.
PF 11-APR-1997; U06049.
PR 12-APR-1996; US-015307.
PA (UYJO) UNIV LELAND STANFORD JUNIOR.
PI Andrew D, Bhanot P, Brink M, Hsieh J, Nathans J,
PI Nusse R, Samos CH, Wangy;
DR N-PSDB; T89882.
PT Identification of Wnt receptor binding modulators - useful for
PT treatment of cancer and growth, development or proliferation related
PT disorders
PS Disclosure: Page 48-50; 61pp; English.
CC This protein comprises the mouse transmembrane receptor,
CC frizzled-8 (Mfz8), encoded by the Mfz8 gene (see T89892). It is
CC an example of a Wnt receptor. Other novel frizzled family members
CC have been identified in human, mouse and Caenorhabditis elegans
CC (see W31268-74) and are considered also to be Wnt receptors. Wnt
CC receptors can be used in a novel, claimed method of screening for
CC compounds which modulate the binding of a Wnt polypeptide (secreted
CC proteins involved in cell-to-cell signalling) to a Wnt receptor.
CC Wnt is involved in (mammary) cancer and other processes involving
CC growth, development and proliferation (both normal and abnormal).
CC Modulators identified by the claimed method are useful for
CC treatment of diseases related to these conditions.
SQ Sequence 685 AA;

Query Match 19.0%; Score 394; DB 27; Length 685;
Best Local Similarity 43.8%; Pred. No. 3.38e-28;
Matches 56; Conservative 23; Mismatches 43; Indels 6; Gaps 6;

Db 40 vplckgigynytmnpnfndtdeaglevhgf-piveiqspdkfflcsmytpicle 98
QY 27 LQCHGIEYQNMRLPNLLGHETMKVEVLEQAGAWIPLVMKQCHPDTKKFLCSLFAFVC 85

Db 99 dykklppcrsvcerakagcaplmrgygfawpdmrdrilpegnpdtlcmdynrtd-1t 157
QY 86 DLDETIQCHSLCVQVKDRCAPVMSAFGFPWDMLECDRFP-ODN-D-LCIPLASSDHL 142

Db 158 taapsppr 165
QY 143 PATEEAPK 150

RESULT 3
ID W31271 standard; Protein: 585 AA.
AC W31271:
DT 27-APR-1998 (first entry)
DE Human frizzled-5 protein Mfz5 (Wnt receptor).
KW Wnt receptor; human frizzled-5 protein; Hfz5 gene;
KW signal transduction; cancer; cell growth; cell proliferation.
OS Homo sapiens.
PN WO9739357-A1.
PD 23-OCT-1997.
PF 11-APR-1997; U06049.
PR 12-APR-1996; US-015307.
PA (UYJO) UNIV LELAND STANFORD JUNIOR.

PA (STRD) UNIV LELAND STANFORD JUNIOR.
PI Andrew D, Bhanot P, Brink M, Hsieh J, Nathans J,
PI Nusse R, Samos CH, Wangy;
DR N-PSDB; T89889.
PT Identification of Wnt receptor binding modulators - useful for
PT treatment of cancer and growth, development or proliferation related
PT disorders
PS Disclosure: Page 37-39; 61pp; English.
CC This protein comprises the human transmembrane receptor,
CC frizzled-5 (Hfz5), encoded by the Hfz5 gene (see T89889). It is
CC an example of a Wnt receptor. Other novel frizzled family members
CC have been identified in human, mouse and Caenorhabditis elegans
CC (see W31268-74) and are considered also to be Wnt receptors. Wnt
CC receptors can be used in a novel, claimed method of screening for
CC compounds which modulate the binding of a Wnt polypeptide (secreted
CC proteins involved in cell-to-cell signalling) to a Wnt receptor.
CC Wnt is involved in (mammary) cancer and other processes involving
CC growth, development and proliferation (both normal and abnormal).
CC Modulators identified by the claimed method are useful for
CC treatment of diseases related to these conditions.
SQ Sequence 585 AA;

Query Match 18.0%; Score 374; DB 27; Length 585;
Best Local Similarity 38.8%; Pred. No. 3.66e-26;
Matches 47; Conservative 29; Mismatches 41; Indels 4; Gaps 4;

Db 20 lyvraaasakpvcqei-t-vpmcrgigynlthmpnqfnhdtddeaglevhgf-plvei 76
QY 7 LFGQPDFSVKRSNCKPIPANLQCHGIEYQNMRLPNLLGHETMKVEVLEQAGAWIPLVMK 65

Db 77 qespdlrfllctmytpicldpdyhklppcrsvcerakagcspmlmrgygfawpdmrscdlr 136
QY 66 QCHPDRKFLCSLFAFVCVCLDLETIQCHSLCVQVKDRCAPVMSAFGFPWDMLECDRF 125

Db 137 p 137
QY 126 p 126

RESULT 4
ID W31270 standard; Protein: 537 AA.
AC W31270:
DT 27-APR-1998 (first entry)
DE Mouse frizzled-4 protein Mfz4 (Wnt receptor).
KW Wnt receptor; mouse frizzled-4 protein; Mfz4 gene;
KW signal transduction; cancer; cell growth; cell proliferation.
OS Mus musculus.
PN WO9739357-A1.
PD 23-OCT-1997.
PF 11-APR-1997; U06049.
PR 12-APR-1996; US-015307.
PA (UYJO) UNIV LELAND STANFORD JUNIOR.
PI Andrew D, Bhanot P, Brink M, Hsieh J, Nathans J,
PI Nusse R, Samos CH, Wangy;
DR N-PSDB; T89888.
PT Identification of Wnt receptor binding modulators - useful for
PT treatment of cancer and growth, development or proliferation related
PT disorders
PS Disclosure: Page 34-35; 61pp; English.
CC This protein comprises the mouse transmembrane receptor,
CC frizzled-4 (Mfz4), encoded by the Mfz4 gene (see T89888). It is
CC an example of a Wnt receptor. Other novel frizzled family members
CC have been identified in human, mouse and Caenorhabditis elegans
CC (see W31268-74) and are considered also to be Wnt receptors. Wnt
CC receptors can be used in a novel, claimed method of screening for
CC compounds which modulate the binding of a Wnt polypeptide (secreted
CC proteins involved in cell-to-cell signalling) to a Wnt receptor.
CC Wnt is involved in (mammary) cancer and other processes involving
CC growth, development and proliferation (both normal and abnormal).
CC Modulators identified by the claimed method are useful for

CC treatment of diseases related to these conditions.
SQ Sequence 537 AA;

Query Match 17.8%; Score 369; DB 27; Length 537;
Best Local Similarity 38.3%; Pred. No. 1.18e-25;

Matches 41; Conservative 26; Mismatches 40; Indels 0; Gaps 0;

Db 47 piriamcnglvntkmpnlvghelqtdeqlttftliqvcssqlgfflcsyvvpmc 106

QY 24 PANLQCHGIEYQNMRLPNLLGHETMKVELEQAGAWIPLVMKQCHPDTKKFLCSLFAPVC 83

Db 107 tekinipgpcgmclsvkrcepvrlrefgawpdtlnckspqnd 153

QY 84 LDDLDETIQCHSLCVQVDRCAPVMSAFGFPWPMLECDRFPQDND 130

RESULT 5

ID W31269 standard; Protein: 525 AA.

AC W31269;

DE 27-APR-1998 (first entry)

DE Caenorhabditis frizzled-1 protein Cfz1 (Wnt receptor).

KW Wnt receptor; Caenorhabditis frizzled-1 protein; Cfz1 gene;

KW signal transduction; cancer; cell growth; cell proliferation.

OS Caenorhabditis elegans.

PN WO9739357-A1.

PD 23-OCT-1997.

PF 11-APR-1997; U06049.

PR 12-APR-1996; US-015307.

PA (UYJO) UNIV JOHNS HOPKINS.

PA (STRD) UNIV LELAND STANFORD JUNIOR.

PI Andrew D, Bhanot P, Brink M, Hsieh J, Nathans J,

PI Nusse R, Samos CH, Wangy;

DR WPI: 97-526631/48.

DR N-PSDB: T89887.

PT Identification of Wnt receptor binding modulators - useful for

PT treatment of cancer and growth, development or proliferation related

PT disorders

PS Disclosure: Page 30-31; 61pp; English.

CC This protein comprises the Caenorhabditis putative transmembrane

CC receptor, frizzled-1, encoded by the Cfz1 gene (see T89887). It is

CC an example of a Wnt receptor. Other novel frizzled family members

CC have been identified in human, mouse and Caenorhabditis elegans

CC (see W31268-74) and are considered also to be Wnt receptors. Wnt

CC compounds can be used in a novel, claimed method of screening for

CC proteins involved in the binding of a Wnt polypeptide (secreted

CC Wnt is involved in cell-to-cell signalling) to a Wnt receptor.

CC Wnt is involved in (mammary) cancer and other processes involving

CC growth, development and proliferation (both normal and abnormal).

CC Modulators identified by the claimed method are useful for

CC treatment of diseases related to these conditions.

SQ Sequence 525 AA;

Query Match 17.2%; Score 357; DB 27; Length 525;

Best Local Similarity 31.8%; Pred. No. 1.93e-24;

Matches 54; Conservative 51; Mismatches 57; Indels 8; Gaps 8;

Db 32 encndlpnltfnpnlvdeeswkdsesiltkpllvsvcsqglkfllcsyvfpmcnekl 91

QY 28 QJCHGIEYQNMRLPNLLGHETMKVELEQAGAWIPLVMKQCHPDTKKFLCSLFAPCLDDL 87

Db 92 anpigrpcrlsvgeklpvlsvsfqkwpdvircdkfpennrekmmnk-gpnegg-ai 149

QY 88 DETIQCHSLCVQVDRCAPVMSAFGFPWPMLECDRFP-QDN-D-LCIPLASSDHLPLA 144

Db 150 qderakf-aa-kessedgdndrvediqrevdrlnqkpcqdevflnrskcv 197

QY 145 TEAPKVCCEACKNDDNDINETLCKNDFALKIKV-KEITVINRDTKII 193

RESULT 6

ID W31267 standard; Protein: 694 AA.

AC W31267;

DE 27-APR-1998 (first entry)

DE Drosophila frizzled-2 protein (Wnt receptor).
KW Wnt receptor; Drosophila frizzled-2 protein; Dfz2 gene;
KW wingless receptor; Wg receptor; signal transduction; cancer;
KW cell growth; cell proliferation.
OS Drosophila melanogaster.

FH Key Location/Qualifiers

FT Misc_difference 268

FT /note= "encoded by CAC"

FT Misc_difference 269

FT /note= "encoded by TGG"

FT Misc_difference 348

FT /note= "encoded by TA (apparent 1 nucleotide deletion of codon)"

FT Misc_difference 488

FT /note= "encoded by TTA"

FT Misc_difference 632

FT /note= "encoded by CTG"

FT Misc_difference 633

FT /note= "encoded by GCG"

FT Misc_difference 671

FT /note= "encoded by CG (apparent 1 nucleotide deletion of codon)"

PN WO9739357-A1.

PD 23-OCT-1997.

PF 11-APR-1997; U06049.

PR 12-APR-1996; US-015307.

PA (UYJO) UNIV JOHNS HOPKINS.

PA (STRD) UNIV LELAND STANFORD JUNIOR.

PI Andrew D, Bhanot P, Brink M, Hsieh J, Nathans J,

PI Nusse R, Samos CH, Wangy;

DR WPI: 97-526631/48.

DR N-PSDB: T89885.

PT Identification of Wnt receptor binding modulators - useful for

PT treatment of cancer and growth, development or proliferation related

PT disorders

PS Claim 2; Page 23-25; 61pp; English.

CC This protein comprises the Drosophila frizzled-2 protein encoded by

CC the Dfz2 gene (see T89885). It is a receptor for wingless (Wg),

CC acting as a signal transducing molecule, and is an example of a Wnt

CC receptor (WntR). Other novel frizzled family members have been

CC identified in human, mouse and Caenorhabditis elegans (see

CC W31268-74) and are considered also to be Wnt receptors. Wnt

CC receptors can be used in a novel, claimed method of screening for

CC compounds which modulate the binding of a Wnt polypeptide (secreted

CC proteins involved in cell-to-cell signalling) to a Wnt receptor.

CC Wnt is involved in (mammary) cancer and other processes involving

CC growth, development and proliferation (both normal and abnormal).

CC Modulators identified by the claimed method are useful for

CC treatment of diseases related to these conditions.

SQ Sequence 694 AA;

Query Match 17.1%; Score 355; DB 27; Length 694;

Best Local Similarity 36.2%; Pred. No. 3.07e-24;

Matches 47; Conservative 32; Mismatches 45; Indels 6; Gaps 6;

Db 51 ygvpaip-kdpnlrceetipmcrigvntsfpmhmetqdeaglvghfw-plveik 108

QY 8 FGQPDFSVKRSCKRPIPANLQCHGIEYQNMRLPNLLGHETMKVEY-LEQAGAWIPLVMKQ 66

Db 109 espdlkfllcsmytpicledvnhkplpcvscvcrarsgcapiqqysfswpermacchlp 168

QY 67 CHPDTRKFLCSLFAPVCLDDLDETIQCHSLCVQVDRCAPVMSAFGFPWPMLECDRFP 126

Db 169 lhgdpdnlcm 178

QY 127 -Q-D-NDLCI 133

RESULT 7

ID W31268 standard; Protein: 666 AA.

AC W31268;

DE 27-APR-1998 (first entry)

DE Mouse frizzled-3 protein Mfz3 (Wnt receptor).

KW Wnt receptor; mouse frizzled-3 protein; Mfz3 gene;
 KW signal transduction; cancer; cell growth; cell proliferation.
 OS Mus musculus.
 PN W09739357-A1.
 PD 23-OCT-1997. U06049.
 PF 11-APR-1997; US-015307.
 PR 12-APR-1996; PR-015307.
 PA (UJJO) UNIV JOHNS HOPKINS.
 PA (SPRD) UNIV LELAND STANFORD JUNIOR.
 PI Andrew D, Bhanot P, Brink M, Hsieh J, Nathans J,
 PI Nusse R, Samos CH, Wangy;
 DR WPI: 97-526631/48.
 DR N-PSDB: T89886.
 PT Identification of Wnt receptor binding modulators - useful for
 PT treatment of cancer and growth, development or proliferation related
 PT disorders
 PS Disclosure; Page 27-29; 61pp; English.
 CC This protein comprises the mouse frizzled-3 protein encoded by
 CC the Mfz3 gene (see T89886), and is a Wnt receptor. Other novel
 CC frizzled family members have been identified in human, mouse and
 CC Caenorhabditis elegans (see W31268-74) and are considered also to
 CC be Wnt receptors. Wnt receptors can be used in a novel, claimed
 CC method of screening for compounds which modulate the binding of a
 CC Wnt polypeptide (secreted proteins involved in cell-to-cell
 CC signalling) to a Wnt receptor. Wnt is involved in (mammary) cancer
 CC and other processes involving growth, development and proliferation
 CC (both normal and abnormal). Modulators identified by the claimed
 CC method are useful for treatment of diseases related to these
 CC conditions.
 CC Sequence 666 AA;
 Query Match 15.3%; Score 317; DB 27; Length 666;
 Best Local Similarity 33.3%; Pred. No. 2,03e-20;
 Matches 44; Conservative 29; Mismatches 54; Indels 5; Gaps 5;
 Db 27 scepi-t-lrmcdlpnyattfmpnlhnydqtaalam-pfhpmvnlcdsrdfrflca 83
 QY 19 NCKPIPANLQCHGIEYQNMRLNGLH-ETMKEVLEQAGAWIPLVMKQHPDPTKFLCS 77
 Db 84 lyapicme-ygrvlpccrrlccqaysecsklmemfgvwpmedmecsfpdcdepyrlvd 142
 QY 78 LFAPVCLDDLETTIQCHSLCVQVDRCAPVMSAFGFPWPMLECDRFPQDNDLCIPLAS 137
 Db 143 lnygdptegap 154
 QY 138 SDHLLPAITEAP 149
 RESULT 8
 ID W31272 standard; Protein; 709 AA.
 AC W31272;
 DT 27-APR-1998 (first entry)
 DE Mouse frizzled-6 protein Mfz6 (Wnt receptor).
 KW Wnt receptor; mouse frizzled-6 protein; Mfz6 gene;
 KW signal transduction; cancer; cell growth; cell proliferation.
 OS Mus musculus.
 PN W09739357-A1.
 PD 23-OCT-1997. U06049.
 PF 11-APR-1997; US-015307.
 PR 12-APR-1996; US-015307.
 PA (UJJO) UNIV JOHNS HOPKINS.
 PA (SPRD) UNIV LELAND STANFORD JUNIOR.
 PI Andrew D, Bhanot P, Brink M, Hsieh J, Nathans J,
 PI Nusse R, Samos CH, Wangy;
 DR WPI: 97-526631/48.
 DR N-PSDB: T89890.
 PT Identification of Wnt receptor binding modulators - useful for
 PT treatment of cancer and growth, development or proliferation related
 PT disorders
 PS Disclosure; Page 41-42; 61pp; English.
 CC This protein comprises the mouse putative transmembrane receptor,
 CC frizzled-6 (Mfz6), encoded by the Mfz6 gene (see T89890). It is
 CC an example of a Wnt receptor. Other novel frizzled family members

CC have been identified in human, mouse and Caenorhabditis elegans
 CC (see W31268-74) and are considered also to be Wnt receptors. Wnt
 CC receptors can be used in a novel, claimed method of screening for
 CC compounds which modulate the binding of a Wnt polypeptide (secreted
 CC proteins involved in cell-to-cell signalling) to a Wnt receptor.
 CC Wnt is involved in (mammary) cancer and other processes involving
 CC growth, development and proliferation (both normal and abnormal).
 CC Modulators identified by the claimed method are useful for
 CC treatment of diseases related to these conditions.
 CC Sequence 709 AA;
 Query Match 13.7%; Score 284; DB 27; Length 709;
 Best Local Similarity 33.6%; Pred. No. 3.89e-17;
 Matches 41; Conservative 26; Mismatches 49; Indels 6; Gaps 6;
 Db 24 cepitvp-r-cmkwty-nmtffnlnghydggaavemghflhlanlecsnlemfclca 80
 QY 20 CKPIPANLQCHGIEYQNMRL-ENLLGHETMKVLEQAGAWIPLVMKQHPDPTKFLCSL 78
 Db 81 fipcteqih-vvlprcklcekivsdcklmdtfgirwpeelecnrlphcdtd-vpvtsh 138
 QY 79 FAPVCLDDLETTIQCHSLCVQVDRCAPVMSAFGFPWPMLECDRFPQDNDLCIPLAS 138
 Db 139 ph 140
 QY 139 DH 140
 RESULT 9
 ID R74187 standard; Protein; 581 AA.
 AC R74187;
 DT 24-DEC-1995 (first entry)
 DE Chick p75.
 KW Neural axon out-growth modulator; epidermal growth factor; EGF;
 KW netrin-2; p75; neurodegenerative disease; transgenic animal;
 KW gene therapy.
 OS Gallus sp.
 FH Key
 FT peptide
 FT 1..15
 FT /label= Sig_peptide
 FT 16..260
 FT /label= Domain-VI
 FT /note= "Domain VI is probably involved in protein-
 FT protein complex formation"
 FT 261..429
 FT /label= Domain-V
 FT /note= "Domain V includes EGF-like repeats"
 FT 430..581
 FT /note= "c-terminal region"
 FT W09513367-A1.
 PN 18-MAY-1995.
 PD 08-NOV-1994; U12913.
 PR 12-NOV-1993; US-132019.
 PA (REGC) UNIV CALIFORNIA.
 PA (UYCO) UNIV COLUMBIA NEW YORK.
 PI Dodd J, Jessell T, Kennedy T, Placzek M, Serafini T;
 PI Tessier-Lavigne M;
 DR WPI: 95-194086/25.
 DR N-PSDB: Q92367.
 PT Neural axon out-growth modulators derived from EGF-like repeats of
 PT netrin 1 or netrin 2 - comprise peptide(s) capable of selectively
 PT increasing spinal axon out-growth or directing axon orientation
 PS Claim 1; Page 48-50; 58pp; English.
 CC An E10 chick brain cDNA library was screened with probes based on
 CC netrin-1 (p78) or netrin-2 (p75) sequences to isolate chick p78
 CC and p75 partial cDNA clones. Full-length clones (given in
 CC Q92366-67, respectively) were subsequently obt'd. by 3'RACE. cDNA
 CC is expressed e.g. in COS or insect cells for recombinant p78 and
 CC p75 prodn., used to breed transgenic animals, or for gene therapy.
 CC Sequence 581 AA;
 Query Match 7.1%; Score 148; DB 14; Length 581;
 Best Local Similarity 27.0%; Pred. No. 3.20e-04;

```

RESULT 11
ID W35946 standard; Protein; 604 AA.
AC W35946;
DE 11-MAY-1998 (first entry)
DE Human netrin-1
KW Netrin-1; neuron; growth; differentiation; morphology;
KW neural disease; diagnosis; therapy; drug screening; human.
OS Homo sapiens.
PN WO9740064-A1.
PD 30-OCT-1997.
PF 19-APR-1997; U06452.
PR 13-APR-1996; US-635137.
PA (EXEL-) EXELIXIS PHARM.
PI (REGC) UNIV CALIFORNIA.
PI Kennedy T, Leonardo D, Serafini T, Shyian A, Swimmer C,
PI Tessier-Lavigne M, Zhang Y;
PI WPI: 97-535773/49.
DR N-PSDB; T97129.
PT Human netrin-1 protein and related nucleic acids - useful in
PT modulating neuron growth and screening for compounds for diagnosis
PT or treatment of diseases associated with undesirable growth
PS Claim 1; Page 13-15; 22pp; English.
CC This protein comprises human netrin-1 (see W35946), a protein
CC which is involved in neural axon guidance, and which is especially
CC useful in modulating neural axon outgrowth. Its amino acid
CC sequence was deduced from a cDNA clone (see T97129) isolated from a
CC human foetal brain cDNA library. Neuron growth, differentiation or
CC morphology can be altered by contact with netrin (claimed).
CC Isolated netrin-1 can also be used to screen chemical libraries
CC for candidate drugs suitable for diagnosis or treatment of diseases
CC associated with undesirable neural cell growth, by comparing
CC binding to a netrin binding target with and without the presence of
CC a prospective agent. Agents that modulate the interaction may be
CC useful as pharmaceutical lead compounds (claimed).
SQ Sequence 604 AA;

Query Match 5.8%; Score 121; DB 28; Length 604;
Best Local Similarity 24.8%; Pred. No. 7.10e-02;
Matches 29; Conservative 30; Mismatches 49; Indels 9; Gaps

Db 454 ipvappttaassvee-pedcdsycaskgklknmkkyckkdvavqhlkdkagdw-w 511
QY 133 IPLASSDHLHPATEAPKVGCEA-CKNKDDNDIMETLCKNDFALKIKV-KEITYINRDT 190
Db 512 kftvnliis--vykggtsrtrgdgs-lwirsdlackcpklpkkk-ylilgnaeds 564
QY 191 KILLETQSKTIYKLVNGVSRDLKKSVLWLKD-SLOCTCEMNDINAPLYMGOKGG 246

RESULT 12
ID R74188 standard; Protein; 529 AA.
AC R74188;
DE 24-DEC-1995 (first entry)
DE Mouse p78.
KW Neural axon out-growth modulator; epidermal growth factor; EGF;
KW netrin-1; p78; neurodegenerative disease; transgenic animal;
KW gene therapy.
KW Mus sp.
FH Key
FH misc_difference 296..297
FT FT /note= "unidentified amino acids"
FT FT /note= "unidentified amino acids"
FT FT /note= "unidentified amino acids"
FT FT /note= "unidentified amino acids"
FT FT /note= "unidentified amino acid"
FT FT /note= "unidentified amino acid"
FT FT /note= "unidentified amino acid"
FT FT /note= "unidentified amino acids"
FT FT /note= "unidentified amino acids"
FT FT /note= "unidentified amino acid"

```

Best Local Similarity 30.5%; Pred.No. 3.78e+00;
Matches 25; Conservative 15; Mismatches 36; Indels 6; Gaps 5;

D_b 158 ipvldikqcpemwkaiacslgkpcshksrsgsil-cksdceveillkkcgd-gnkfpedht 215
||| :|: | |: ||| :|: | |: ||| :|: | |: ||| :|:
Q_y 60 IP-LVNMKQHPT-KKFLCSLFPVCVLDDLETIOCHSLCVQVKDRCAPYMSAFGFWP 117
||| :|: | |: ||| :|: | |: ||| :|: | |: ||| :|:

D_b 216 aesicellstddlknclptdt 237
||| :|: | |: ||| :|: | |: ||| :|: | |: ||| :|:
Q_y 118 DMLECDRFFQNDL--CIPLAS 137
||| :|: | |: ||| :|: | |: ||| :|: | |: ||| :|:

RESULT 14

ID	W53368 standard; Protein; 804 AA.
AC	W53368;
DE	17-JUN-1998 (first entry)
DE	H. pylori ORF 09apll406orf2 protein.
KW	Cytoplasmic; vaccine; prevention; treatment; infection; envelope;
KW	identification; binding compound; bacteria; life cycle; activator;
KW	inhibitor; duodenal ulcer disease; chronic gastritis; diagnosis.
OS	Helicobacter Pylori.
PN	WO9737044-Al.
PD	09-OCT-1997.
PR	27-MAR-1997; U05223.
PR	06-DEC-1996; US-761318.
PR	29-MAR-1996; US-625811.
PR	02-APR-1996; US-758731.
PR	25-OCT-1996; US-736905.
PR	28-OCT-1996; US-738859.
PA	(ASTR) ASTRA AB.
PI	Alm RA, Smith D;
DR	WPI; 97-503122/46.
DR	N-PSDB: V24777.
CC	Helicobacter pylori nucleic acid sequences and encoded
PT	polypeptide(s) - useful in vaccines to treat or prevent H. pylori
PT	infection and for diagnosis of H. pylori infection
PS	Claim 14; Page 587-589; 1145pp; English.
CC	This sequence is a H. pylori protein of unspecified function.
CC	The protein may be used in a vaccine to prevent or treat H. pylori
CC	infection or to identify H. pylori polypeptide binding compounds,
CC	useful as potential H. pylori life cycle activators or inhibitors. The
CC	DNA and probes derived from it may be used for the identification of
CC	H. pylori in a sample and the diagnosis of H. pylori infection. Nucleic
CC	acid sequences complementary to the DNA act as antisense sequences and
CC	can be used to prevent the translation of H. pylori mRNA. Antibodies
CC	against the protein can be used in immunoassays to evaluate the abundance
CC	and distribution of H. pylori-specific antigens. The genomic sequence of
CC	H. pylori (ATCC 55679) was determined from overlapping contigs generated
CC	by mechanically shearing the bacterial DNA. The sequences were analysed
CC	for ORF of at least 180 nucleotides, and the predicted coding regions
CC	defined by computer evaluation. To identify likely H. pylori antigens for
CC	vaccine development, the amino acid sequences predicted from various ORF
CC	were analysed for significant homology to other known or exported
CC	membrane proteins. Having identified and determined the sequences of
CC	interest, particular regions can be isolated from H. pylori by PCR
CC	amplification for recombinant polypeptide production, e.g. in E. coli
CC	hosts.
SQ	Sequence 804 AA;

Query Match 4.8%; Score 99; DB 29; Length 804;
Best Local Similarity 31.5%; Pred.No. 4.53e+00;
Matches 17; Conservative 16; Mismatches 19; Indels 2; Gaps 2;

D_b 78 vdisipkeftplenaffnitymrqqgfihlpkvmrrltlsfdrydkkal 131
::::|: |: | ::||| :|: | |: ||| :|: | |: ||| :|:
Q_y 165 IMETLKNDFA-LKIKVEITYINRDTRIIKETSKTIYLNGVS-BRDJUKSV 216
::::|: |: | ::||| :|: | |: ||| :|: | |: ||| :|:

RESULT 15

ID	R44929 standard; Protein; 15281 AA.
AC	R44929;
DE	08-JUL-1994 (first entry)
DE	T. niveum Cyclosporin synthetase.

KW Enzyme; cyclosporin; synthetase-like activity; Tolypocladium niveum;
 KW T. inflatum GAMS; biosynthesis; vector; cyclosporin synthetase.

OS Tolypocladium niveum.

PN EP-578616-A.

PD 12-JAN-1994.

PF 05-JUL-1993; 810474.

PR 09-JUL-1992; AT-001403.

PR 08-MAR-1993; AT-000437.

PR 29-APR-1993; CH-001310.

PR 04-MAY-1993; CH-001375.

PA (SANO) SANDOZ LTD.

PA (SANO) SANDOZ PATENT GMBH.

PA (SANO) SANDOZ-ERFINDUNGEN VERW GES MBH.

PI Leitner E, Schneider E, Schoengendorfer K, Weber G;

DR WPI: 94-010432/02.

DR N-PSDB: Q54385.

PT Isolated DNA sequence - which codes for enzyme having cyclosporin

PT synthetase like activity

PS Claim 1; Page 41-84; 93pp; English.

CC This sequence represents an enzyme which has cyclosporin synthetase-

CC like activity. This sequence was isolated from Tolypocladium niveum

CC (formerly known as T. inflatum GAMS). This enzyme catalyses the

CC peptide biosynthesis of cyclosporins and structurally related

CC molecules. This sequence may be used for the production of

CC cyclosporin by transforming a vector containing this sequence in

CC to a recombinant host. This allows effective production of anti-

CC biotic cyclosporin or its derivatives.

CC Sequence 15281 AA;

SQ

Query Match 4.7%; Score 97; DB 9; Length 15281;

Best Local Similarity 26.7%; Pred. No. 6.52e+00;

Matches 24; Conservative 23; Mismatches 35; Indels 8; Gaps 8;

Db 3819 lrgtkisdhiainipnsktiverticesvydlgg-dakdsndrswlsaa-rsnakva 3876

QY 176 LK-IKVKE-ITVIN-RDTRILE-TKSKTYIKLVGYSERDLKSLWLKDSLQCTCEMN 231

Db 3877 slsaldldiagaeagfrveiscarqwsqg 3906

QY 232 DINAPYLW-MGOKQGGELVITSVKRW-QKG 259

Search completed: Tue Aug 4 10:40:20 1998

Job time : 71 secs.

THIS PAGE BLANK (use pto)


```
ACCESSIONS S78444
REFERENCE S78444
#authors Bhanot, P.; Wang, Y.; Nathans, J.
#journal submitted to the EMBL Data Library, July 1996
#accession S78444
#status preliminary
#molecule_type DNA
#residues 1-694 #label BHA
#cross-references EMBL:U65589
#length 694 #molecular-weight 75437 #checksum 3113
SUMMARY

Query Match 17.1%; Score 355; DB 2; Length 694;
Best Local Similarity 36.2%; Pred. No. 1.19e-45;
Matches 47; Conservative 32; Mismatches 45; Indels 6; Gaps 6;

Db 51 YGPAIP-KDNLRCETITPMCRGIGYNTSPFNEMNHETQDEAGLEVHGFQFW-PLVEIK 108
QY 8 FGQPDFSYKRSNCKRPIPANLQCHGIEYQNMRLPNLGLHETMKREV-LEQAGAWIPLVMKQ 66
Db 109 CSPDLKFFLCSMTPIGLEDYHKPLPVCRSVCERARSGCAPIMQOYSFEWPERMACEHLP 168
QY 67 CHPTDKFLCSLFAVPVCLDDLETIQCHSLCVQKDRCAPVMSAFGFPWPDMLECDRFP 126
Db 169 LHGDPDNLCM 178
QY 127 -Q-D-NDLCI 133

RESULT 3
ENTRY S71786 #type complete
TITLE dfz2 protein - fruit fly (Drosophila melanogaster)
ORGANISM Drosophila melanogaster
DATE 12-Feb-1998 #sequence_revision 13-Mar-1998 #text_change
13-Mar-1998
ACCESSIONS S71786
REFERENCE S71786
#authors Bhanot, P.; Brink, M.; Harryman Samos, C.; Hsieh, J.C.; Wang,
Y.; Macke, J.P.; Andrew, D.; Nathans, J.; Nusse, R.
#journal Nature (1996) 382:225-230
#title A new member of the frizzled family from Drosophila functions
as a wingless receptor.
#accession S71786
#status preliminary; nucleic acid sequence not shown
#molecule_type DNA
#residues 1-694 #label BHA
#cross-references EMBL:U65589
#length 694 #molecular-weight 75423 #checksum 3095
SUMMARY

Query Match 17.1%; Score 355; DB 2; Length 694;
Best Local Similarity 36.2%; Pred. No. 1.19e-45;
Matches 47; Conservative 32; Mismatches 45; Indels 6; Gaps 6;

Db 51 YGPAIP-KDNLRCETITPMCRGIGYNTSPFNEMNHETQDEAGLEVHGFQFW-PLVEIK 108
QY 8 FGQPDFSYKRSNCKRPIPANLQCHGIEYQNMRLPNLGLHETMKREV-LEQAGAWIPLVMKQ 66
Db 109 CSPDLKFFLCSMTPIGLEDYHKPLPVCRSVCERARSGCAPIMQOYSFEWPERMACEHLP 168
QY 67 CHPTDKFLCSLFAVPVCLDDLETIQCHSLCVQKDRCAPVMSAFGFPWPDMLECDRFP 126
Db 169 LHGDPDNLCM 178
QY 127 -Q-D-NDLCI 133

RESULT 4
ENTRY S15709 #type complete
TITLE hypothetical protein - fruit fly (Drosophila melanogaster)
ORGANISM Drosophila melanogaster
DATE 21-Nov-1993 #sequence_revision 10-Nov-1995 #text_change
16-Feb-1997
ACCESSIONS S15709
REFERENCE S15709
```

```
#authors Adler, P.N.; Vinson, C.; Park, W.J.; Conover, S.; Klein, L.
#journal Genetics (1990) 126:401-416
#title Molecular structure of frizzled, a Drosophila tissue polarity
gene.
#cross-references MUID:91060073
#accession S15709
#status preliminary
#molecule_type DNA
#residues 1-415 #label ADL
#cross-references EMBL:X54648
GENETICS
#gene FlyBase:fz
#cross-references FlyBase:FBgn0001085
#introns 224/3; 264/3; 329/3; 405/1
#length 415 #molecular-weight 46117 #checksum 4401
SUMMARY

Query Match 16.0%; Score 332; DB 2; Length 415;
Best Local Similarity 37.5%; Pred. No. 3.34e-41;
Matches 45; Conservative 22; Mismatches 47; Indels 6; Gaps 6;

Db 50 HNRCEPI-T-ISICKNIPYNTIMPNLIGHTKQEEAGLEVHGFAPLVKIGSCDDLLQLFLC 107
QY 17 RSNCKRPIANLQCHGIEYQNMRLPNLGLHETMKREVLEQAGAWIPLVMKQCHPDTKKFLC 76
Db 108 SLYVPVC-TILERPIPPCRSLCESARV-CEKLMKTYNFNPNLENCSKFPVHGEGDLQVA 165
QY 77 SLFAPVCLDDLETIQCHSLCVQKDRCAPVMSAFGFPWPDMLECDRFP-QD-NDLCIP 134

RESULT 5
ENTRY S03540 #type complete
TITLE gene frizzled protein precursor - fruit fly (Drosophila
melanogaster)
ORGANISM Drosophila melanogaster
DATE 31-Mar-1990 #sequence_revision 31-Mar-1990 #text_change
23-Aug-1997
ACCESSIONS S03540; S15708
REFERENCE S03540
#authors Vinson, C.R.; Conover, S.; Adler, P.N.
#journal Nature (1989) 338:263-264
#title A Drosophila tissue polarity locus encodes a protein
containing seven potential transmembrane domains.
#cross-references MUID:89159415
#accession S03540
#status not compared with conceptual translation
#molecule_type mRNA
#residues 1-581 #label VIN
REFERENCE S15708
#authors Adler, P.N.; Vinson, C.; Park, W.J.; Conover, S.; Klein, L.
#journal Genetics (1990) 126:401-416
#title Molecular structure of frizzled, a Drosophila tissue polarity
gene.
#cross-references MUID:91060073
#accession S15708
#status preliminary
#molecule_type DNA
#residues 1-581 #label ADL
#cross-references EMBL:X54648
GENETICS
#gene FlyBase:fz
#cross-references FlyBase:FBgn0001085
#introns 224/3; 264/3; 329/3; 405/1
#length 581 #molecular-weight 64846 #checksum 358
#product gene frizzled protein #status predicted #label SIG\
WAT
SUMMARY

Query Match 16.0%; Score 332; DB 2; Length 581;
Best Local Similarity 37.5%; Pred. No. 3.34e-41;
Matches 45; Conservative 22; Mismatches 47; Indels 6; Gaps 6;
```

QY 133 IPLASSDHLLPATEAPKVCEA-CKNNKDDNDINMETLCKNDFALIKVKETIYNRDTK 191
Dd 490 FTNIILS--VYKCORDERVRK-GDNFLWIHLKLDSCKPKI-QISKKYLVNG 536
 :::: | ::| : ||: | -|- | -|:- |: |::::
QY 192 ILLETSKTYLYKNVSERDKKSJWLK-DSLQCCTEEMNDINAPYLNG 241

RESULT	8
ENTRY	A54665 #type complete
TITLE	netrin-1 precursor - chicken
ORGANISM	#formal_name Gallus gallus #common_name chicken
DATE	06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 10-Sep-1997

#journal J. Neurosci.
#title The netrins define a family of axon outgrowth-promoting proteins homologous to *Drosophila* netrin and *C. elegans* UNC-6.
#author Tessier-Lavigne, M.
#year 1994
#volume 14
#pages 424-438
#publisher Society for Neuroscience
#city Washington, DC
#country USA
#abstract The netrins define a family of axon outgrowth-promoting proteins homologous to *Drosophila* netrin and *C. elegans* UNC-6. In the vertebrate nervous system, netrin is expressed in the floor plate and in the developing brain, and it has been shown to be a potent chemoattractant for reticulospinal axons. We have now identified and characterized the complementary DNAs for the other members of the netrin family. These include two netrin-like proteins, netrin-1 and netrin-2, and a novel protein, netrin-3. Netrin-1 and netrin-2 are expressed in the floor plate and in the developing brain, while netrin-3 is expressed in the developing brain and in the peripheral nervous system. Netrin-1 and netrin-2 are also potent chemoattractants for reticulospinal axons. The netrins define a family of axon outgrowth-promoting proteins homologous to *Drosophila* netrin and *C. elegans* UNC-6. In the vertebrate nervous system, netrin is expressed in the floor plate and in the developing brain, and it has been shown to be a potent chemoattractant for reticulospinal axons. We have now identified and characterized the complementary DNAs for the other members of the netrin family. These include two netrin-like proteins, netrin-1 and netrin-2, and a novel protein, netrin-3. Netrin-1 and netrin-2 are expressed in the floor plate and in the developing brain, while netrin-3 is expressed in the developing brain and in the peripheral nervous system. Netrin-1 and netrin-2 are also potent chemoattractants for reticulospinal axons.

SUMMARY	#length	606	#molecular-weight	68126	#checksum	5400
---------	---------	-----	-------------------	-------	-----------	------

```

QY      133  IPASSDHLLPATEAPKVC EA-CKKNDDDDDDIMETLCKNDFALAKIV-KEITYINRDT 190

Db      514  KFTVNIIS--VYK-QG-SNR-LRRGQOTUWVHAKDIACKCPKVVKMKK-YLLGLGSTE DS 566

QY      191  KIILEYKSTYIKYNGVSRDLKKS--VWLMD-LSQCTCEENDDINAYLVNGORQO 246

```

QY 191 KIILETKSKYIYKLVGVSERDLKKS--VLWLKD-SLQCTCEEMNDINAPILVMGQRQGG 246

RESULT 9

ENTRY	JC5539	#type complete
TITLE	Smoothed protein - mouse	
ORGANISM	#formal_name Mus musculus #common_name house mouse	
DATE	02-Sep-1997 #sequence revision 05-Sep-1997 #text-change	

05-Sep-1997
JC5539; PC4476
JC5539
ACCESSIONS
REFERENCE
Aktiyana, H.; Shigeno, C.; Hiraki, Y.; Shukunami, C.; Kohno,

```

#accession      JC5539
##molecule_type mRNA
##residues      1-793  ##label AKI
#accession      PC4476
##molecule_type protein
##residues      528-533;539-545;600-605  ##label AK2
##experimental_source Atr5 cell
COMMENT
This protein is used in the conserved targets in Hedgehog
signalling pathway, together with patched and Gli. These protein
are responsible for the skeletal abnormalities in Gorlin and
Greig syndromes.
cartilage differentiation in clonal mouse EC cells, Atr5C5.

```

```

1-32 . #domain signal sequence #status predicted #label size
SUMMARY #length 793 #molecular-weight 87299 #checksum 8609

Query Match          5.9%  Score 122:  DB 2:  Length 793;
Best Local Similarity 26.7%;  Pred. No. 3.66e-04;
Matches 31;  Conservative 22; Mismatches 54;  Indels 9;  Gaps

```

[illegible]

```
ENTRY      152313      #type complete
TITLE      cytochrome P450 2B4 homolog, brain (clone 2d-29) - rat
ORGANISM   #formal_name Rattus norvegicus #common_name Norway rat
DATE       27-Feb-1997 #sequence_revision 27-Feb-1997 #text_change
23-Jan-1998

ACCESSIONS 152313
REFERENCE   152313
#authors   Kawashima, H.; Strobel, H.W.
#journal   Biochem. Biophys. Res. Commun. (1995) 209:535-540
#title     cDNA cloning of a novel rat brain cytochrome P450 belonging
           to the CYP2D subfamily.
#cross-references MUID:95251650
#accession 152313
#status    preliminary; translated from GB/EMBL/DBJ
#molecule_type mRNA
#residues  1-500 #label RES
#cross-references GB:S77859; NID:g998524; PID:g998525
#experimental_source brain, strain Sprague-Dawley
CLASSIFICATION #superfamily cytochrome P450
KEYWORDS      heme; transmembrane protein
FEATURE       446
             #binding_site heme iron (Cys) (axial ligand) #status
             predicted
SUMMARY       #length 500 #molecular-weight 56683 #checksum 8540
             4.8%; Score 99; DB 2; Length 724;
             Query Match 4.8%; Score 100; DB 2; Length 500;
             Best Local Similarity 26.7%; Pred. No. 3.97e-01;
             Matches 16; Conservative 22; Mismatches 18; Indels 4; Gaps 3;

Db 173 PFPS-PNTLLKAVCNVIALSLFACR-FEYNDPRFIRLL--DLLKDTLEESGFLPMLLN 228
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 6 FLFGQPDFSYKRSNCKPIPANLQLCHGIEYQNMRLNLLGHETMKVEVLEQAGAWPLVMK 65

RESULT 14
ENTRY   B32571      #type complete
TITLE   ribosomal protein S6 kinase II (EC 2.7.-.-) alpha chain
ORGANISM #homolog (clone Mu6a) - mouse
#formal_name Mus musculus #common_name house mouse
DATE     31-Mar-1990 #sequence_revision 31-Mar-1990 #text_change
26-Feb-1998

ACCESSIONS B32571
REFERENCE   A93113
#authors   Alcorta, D.A.; Crews, C.M.; Sweet, L.J.; Bankston, L.; Jones,
           S.W.; Erikson, R.L.
#journal   Mol. Cell. Biol. (1989) 9:3850-3859
#title     Sequence and expression of chicken and mouse rsk: homologs of
           Xenopus laevis ribosomal S6 kinase.
#cross-references MUID:89384612
#accession B32571
#molecule_type mRNA
#residues  1-724 #label ALC
#cross-references GB:M23489
CLASSIFICATION #superfamily ribosomal protein S6 kinase II; protein kinase
               homology
KEYWORDS      ATP; phosphotransferase; serine/threonine-specific protein
               kinase
FEATURE       60-310
             #domain protein kinase homology #label KIN1\
68-76      #region protein kinase ATP-binding motif\
405-664    #domain protein kinase homology #label KIN2
SUMMARY      #length 724 #molecular-weight 81594 #checksum 4206;

Query Match 4.8%; Score 99; DB 2; Length 724;
Best Local Similarity 35.2%; Pred. No. 5.34e-01;
Matches 19; Conservative 13; Mismatches 18; Indels 4; Gaps 4;

Db 499 KI-LRQKFSEAEAFVLHTISKTYEVLHSQGVHRLKPSNLYVDESGNPEC 541
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 177 KIKVKEITYINRDKIILETKSKI-Y-KLNGVSRDLKKS-VLWIKDSLOQTC 227

RESULT 15
```

```
ENTRY      E64474      #type complete
TITLE      hypothetical protein MJ1398 - Methanococcus jannaschii
ORGANISM   #formal_name Methanococcus jannaschii
DATE       13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change
10-Oct-1997

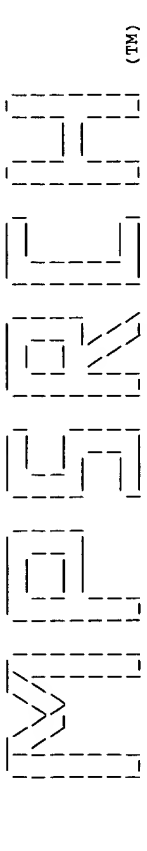
ACCESSIONS E64474
REFERENCE   A64300
#authors   Bult, C.J.; White, O.; Olsen, G.J.; Zhou, L.; Fleischmann,
           R.D.; Sutton, G.G.; Blake, J.A.; Fitzgerald, L.M.; Clayton,
           R.A.; Gocayne, J.D.; Kerlavage, A.R.; Dougherty, B.A.;
           Tomb, J.F.; Adams, M.D.; Reich, C.I.; Overbeek, R.;
           Kirkness, E.F.; Weinstock, K.G.; Merrick, J.M.; Glodek, A.;
           Scott, J.L.; Geoghegan, N.S.M.; Weidman, J.F.; Fuhrmann,
           J.L.; Nguyen, D.; Utterback, T.R.; Kelley, J.M.; Peterson,
           J.D.; Sadow, P.W.; Hanna, M.C.; Cotton, M.D.; Roberts,
           K.M.; Hurst, M.A.; Kaine, B.P.; Borodovsky, M.; Klenk,
           H.P.; Fraser, C.M.; Smith, H.O.; Woese, C.R.; Venter, J.C.
#journal   Science (1996) 273:1058-1073
#title     Complete genome sequence of the methanogenic archaeon,
           Methanococcus jannaschii.
#cross-references MUID:96337999
#accession E64474
#status    preliminary; nucleic acid sequence not shown;
           translation not shown
#molecule_type DNA
#residues  1-391 #label BUL
#cross-references GB:U67580; GB:L77117; NID:g1592044; PID:g1592046;
           TIGR:MJ1398; PID:g1511417

GENETICS
#map_position FOR1358185-1359360
SUMMARY     #length 391 #molecular-weight 45529 #checksum 6984
           4.7%; Score 98; DB 2; Length 391;
           Query Match 4.7%; Score 98; DB 2; Length 391;
           Best Local Similarity 30.0%; Pred. No. 7.17e-01;
           Matches 15; Conservative 19; Mismatches 14; Indels 2; Gaps 2;

Db 84 LRFKIENKSNLIITTKNNVNTNKMIIFTSINGSRVNGNKTIIWIKDPL 133
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 176 LKIKVKEITYINRDKIILETKSKI-YK-LNGVSRDLKKSVLWIKDSL 223

Search completed: Tue Aug 4 10:41:38 1998
Job time : 61 secs.
```

THIS PAGE BLANK (USPTO)



Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (C) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

MPsrch_pp protein - protein database search, using Smith-Waterman algorithm
Run on: Tue Aug 4 10:41:56 1998; MasPar time 10.08 Seconds
Tabular output not generated. 684.576 Million cell updates/sec

Title: >US-08-848-439-3
Description: (1-275) from US08848439.pep
Perfect Score: 2073
Sequence: 1 SARGLFQGPDFSYSRNC.....WQKGRFKEKRSIRKLCQ 275

Scoring table: PAM 150
Gap 11

Searched: 69111 seqs, 25083644 residues

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: swiss-prot35
1:swiss1

Statistics: Mean 46.355; Variance 72.254; scale 0.642

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	ID	Description	Pred. No.
1	332	16.0	FRIZ_DROME	FRIZZLED PROTEIN PRECU	5.63e-52
2	148	7.1	NET1_CHICK	NETRIN-2 PRECURSOR (FR	1.08e-10
3	124	6.0	NET1_CHICK	NETRIN-1 PRECURSOR.	3.47e-06
4	109	5.3	YQ18_CAEEL	HYPOTHETICAL 53.8 KD P	1.39e-03
5	101	4.9	RPCL_GIALA	DNA-DIRECTED RNA POLYM	2.78e-02
6	100	4.8	CPDI_RAT	CYTCHROME P450 IID18	3.99e-02
7	100	4.8	CPDI_RAT	CYTCHROME P450 IID4	3.99e-02
8	99	4.8	FRIZ_MOUSE	RIPOSONAL PROTEIN S6 K	5.72e-02
9	98	4.7	VG07_BPT4	BASEPLATE STRUCTURAL P	8.18e-02
10	96	4.6	ION3_CARAU	INTERMEDIATE FILAMENT	1.66e-01
11	94	4.5	UD13_RAT	UDP-GLUCURONOSYLTRANSF	3.32e-01
12	92	4.4	YMB1_YEAST	HYPOTHETICAL 180.2 KD	6.59e-01
13	90	4.3	GYRB_ACIS3	DNA GYRASE SUBUNIT B	1.29e+00
14	89	4.3	VIIRF_SHIDY	VIRULENCE REGULON TRAN	1.79e+00
15	90	4.3	KSGA_MYCPN	DIMETHYLADEINOSINE TRAN	1.29e+00
16	90	4.3	VATA_MAIZE	VACUOLAR ATP SYNTHASE	1.29e+00
17	89	4.3	KSGA_CHICK	RIPOSONAL PROTEIN S6 K	1.79e+00
18	88	4.2	YB09_YEAST	HYPOTHETICAL 14.8 KD P	2.48e+00
19	88	4.2	YLF2_CAEEL	HYPOTHETICAL 20.1 KD P	2.48e+00
20	88	4.2	AMZ2_ORISA	ALPHA-AMYLASE ISOZYME	2.48e+00
21	88	4.2	AMZ2_ORISA	ALPHA-AMYLASE ISOZYME	2.48e+00
22	87	4.2	YHE2_PSEAE	HYPOTHETICAL 50.0 KD P	3.43e+00
23	87	4.2	NOAL_HUMAN	ONCONEURAL VENTRAL ANT	3.43e+00

24	87	4.2	580	1	VATA_HORVU	VACUOLAR ATP SYNTHASE	3.43e+00
25	88	4.2	633	1	KS62_MOUSE	RIPOSONAL PROTEIN S6 K	2.48e+00
26	88	4.2	740	1	KS62_HUMAN	RIPOSONAL PROTEIN S6 K	2.48e+00
27	88	4.2	848	1	NEM_MOUSE	NEUROFILAMENT TRIPLET	2.48e+00
28	88	4.2	1102	1	YK64_CAEEL	HYPOTHETICAL 124.8 KD	2.48e+00
29	87	4.2	1201	1	COPA_YEAST	COTOMER ALPHA SUBUNIT	3.43e+00
30	87	4.2	1752	1	DESP_HUMAN	DESMOPLAKIN I AND II	3.43e+00
31	87	4.2	2749	1	IP3R_MOUSE	INOSITOL 1,4,5-TRISPHO	3.43e+00
32	87	4.2	2749	1	IP3R_RAT	INOSITOL 1,4,5-TRISPHO	3.43e+00
33	85	4.1	216	1	GYRB_ACIS7	DNA GYRASE SUBUNIT B	6.48e+00
34	86	4.1	327	1	VMP_CAWD	MOVEMENT PROTEIN (CELL	4.72e+00
35	86	4.1	450	1	DCOR_CHICK	ORNITHINE DECARBOXYLAS	4.72e+00
36	86	4.1	461	1	DCOR_BOVIN	ORNITHINE DECARBOXYLAS	4.72e+00
37	85	4.1	495	1	EIBL_ADE02	E1B PROTEIN, LARGE T-A	6.48e+00
38	86	4.1	556	1	FTHS_STRUM	FORMATE--TETRAHYDROFOL	4.72e+00
39	86	4.1	620	1	Y870_METJA	HYPOTHETICAL PROTEIN M	4.72e+00
40	85	4.1	633	1	YHR4_YEAST	HYPOTHETICAL 71.2 KD P	6.48e+00
41	86	4.1	725	1	VACB_MYCGE	VACB PROTEIN HOMOLOG.	4.72e+00
42	85	4.1	798	1	VP16_YEAST	VACUOLAR PROTEIN SORTI	6.48e+00
43	85	4.1	1407	1	CYAA_DICDI	ADENYLATE CYCLASE, AGG	6.48e+00
44	85	4.1	1726	1	MSPI_PLAEP	MEROZOITE SURFACE PROT	6.48e+00
45	85	4.1	2688	1	ZEPI_MOUSE	ZINC FINGER PROTEIN 40	6.48e+00

ALIGNMENTS

RESULT 1
ID FRIZ_DROME STANDARD; PRT; 581 AA.
AC P18537;
DT 01-NOV-1990 (REL. 16, CREATED)
DT 01-NOV-1990 (REL. 16, LAST SEQUENCE UPDATE)
DT 01-MAY-1992 (REL. 22, LAST ANNOTATION UPDATE)
DE FRIZZLED PROTEIN PRECURSOR.
GN FZ
OS DROSOPHILA MELANOGASTER (FRUIT FLY).
OC EUKARYOTA; METAZOA; ARTHROPODA; INSECTA; DIPTERA.
RN [1]
RP SEQUENCE FROM N.A. (CLASS I).
RX MEDLINE; 89159415.
RA VINSON C.R., CONOVER S., ADLER P.N.;
RL NATURE 338:263-264(1989).
[2]
RP SEQUENCE FROM N.A. (CLASSES I AND II).
RX MEDLINE; 91060073.
RA ADLER P.N., VINSON C., PARK W.J., CONOVER S., KLEIN L.;
RL GENETICS 126:401-416(1990).
CC -!- FUNCTION: REQUIRED TO COORDINATE THE CYTOSKELETONS OF EPIDERMAL
CELLS TO PRODUCE A PARALLEL ARRAY OF CUTICULAR HAIRS AND BRISTLES.
CC FZ IS REQUIRED FOR THE TRANSMISSION OF POLARITY INFORMATION, AND
TRANSDUCTION.
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC -!- ALTERNATIVE PRODUCTS: CLASS I AND CLASS II PROTEINS DIFFER AT THE
C-TERMINAL FROM AA 406 DUE TO USE OF ALTERNATIVE 3' EXONS.
DR EMBL; X54648; G804979; JOINED.
DR EMBL; X54649; G804979; JOINED.
DR EMBL; X54650; G804979; JOINED.
DR EMBL; X54651; G804979; JOINED.
DR EMBL; X54648; G804980; JOINED.
DR EMBL; X54649; G804980; JOINED.
DR EMBL; X54650; G804980; JOINED.
DR EMBL; X54652; G804980; JOINED.
DR EMBL; X54646; G7981; JOINED.
DR EMBL; X54647; G7983; ALT_SEQ.
DR PIR; S03540; S03540.
DR PIR; S15708; S15708.
DR FLYBASE; FBgn0001085; fz.
KW TRANSMEMBRANE; SIGNAL; ALTERNATIVE SPLICING; DEVELOPMENTAL PROTEIN.
FT SIGNAL 1 26
FT CHAIN 27 581
FT TRANSMEM 248 270
FT TRANSMEM 281 303
FT TRANSMEM 336 368
FT TRANSMEM 381 401
FT TRANSMEM 381 401

FT	TRANSMEM	422	439	POTENTIAL.
FT	TRANSMEM	472	492	POTENTIAL.
FT	TRANSMEM	530	553	POTENTIAL.
FT	VARSPIC	406	415	DILSGVCFVG -> MYLWQFHTIN (IN CLASS II).
FT	VARSPIC	416	581	MISSING (IN CLASS II).
SQ	SEQUENCE	581 AA:	64847 MW:	279EBE85 CRC32:

```
Query Match          16.0%;      Score 332;  DB 1;  Length 581;
Best Local Similarity 37.5%;      Pred. NO. 5.63e-52;
Matches 45; Conservative 22; Mismatches 47; Indels 6; Gaps 6;
```

Db	50	HNRCEPI-T-ISTCKNPTNWTMPNLGHTKQBEAGLEVHQFAPVAKVGCSDQLQFLC	107
		:::KKK::PPANLQJCHGEYQNMELPNLLGHETKWELEQAGAIPLVMKQCHPDTKKFLC	176
Qy	17	RSNCKP:PPANLQJCHGEYQNMELPNLLGHETKWELEQAGAIPLVMKQCHPDTKKFLC	76
Db	108	SLXVPVC-T-TILERIPPCRSLSCEARSV-CEKLMKTYNFMNPENLECSKFPVHGGE	165
		LVLA::PPANLQJCHGEYQNMELPNLLGHETKWELEQAGAIPLVMKQCHPDTKKFLC	176
Qy	77	SLFAPVCLDDLDDETIOQCHSLVCVQKDRCAFPVMSAEGFPWFDLMCEDRFFQD-NDLCIP	134
		LVLA::PPANLQJCHGEYQNMELPNLLGHETKWELEQAGAIPLVMKQCHPDTKKFLC	176

RESULT	2	STANDARD	PRT	581 AA.
ID	NET2_CHICK			
AC	Q90923;			
DT	01-NOV-1997 (REL. 35, CREATED)			
DT	01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)			
DT	01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)			
DE	NETRIN-2 PRECURSOR (FRAGMENT).			
OS	GALLUS GALLUS (CHICKEN).			
OC	EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; AVES; NEOGNATHAE;			
OC	GALLIFORMES.			

SEQUENCE FROM N.R.A.
NF STRAIN-WHITE LEIGHORN; TISSUE-EMBRYONIC BRAIN;
RX MEDLINE: 94340732.
RA SERAFINI T., KENNEDY T.E., GALKO M.J., MIRZAYAN C., JESSELL T.M.,
RA TESSIER-LAVIGNE M.;
RL CELL 78:409-424(1994).
CC -!- FUNCTION: NETRINS CONTROL GUIDANCE OF CNS COMMISSURAL AXONS AND
CC -!- PERIPHERAL MOTOR AXONS.
CC -!- SUBCELLULAR LOCATION: EXTRACELLULAR.
CC -!- SIMILARITY: CONTAINS ONE LAMININ N-TERMINAL DOMAIN (DOMAIN VI).
CC -!- SIMILARITY: CONTAINS 3 LAMININ EGF-LIKE DOMAINS.
CC -!- SIMILARITY: CONTAINS ONE C345C DOMAIN.
CC -!- SIMILARITY: TO CAENORHARDITIS ELIGANS UNC-6 PROTEIN.

DR PROSITE; PS00022; EGF_1; 2.
DR PROSITE; PS01248; LAMININ_TYPE_EGF; 3.
KW GLYCOPROTEIN; EXTRACELLULAR MATRIX; SIGNAL; LAMININ EGF-LIKE DOMAIN;
KW REPEAT.

FT	NON-TIER	1	
FT	SIGNAL	<1	POTENTIAL.
FT	CHAIN	16	NETRIN-2.
FT	DOMAIN	16	LAMININ N-TERMINAL (DOMAIN VI).
FT	DOMAIN	262	3 X LAMININ EGF-LIKE REPEATS (LAMININ DOMAIN V-LIKE).
FT		430	

FT	DOMAIN	202	LAMININ EGF-LIKE 2.	317	LAMININ EGF-LIKE 1.
FT	DOMAIN	318	LAMININ EGF-LIKE 2.	318	LAMININ EGF-LIKE 2.
FT	DOMAIN	381	LAMININ EGF-LIKE 3.	430	LAMININ EGF-LIKE 3.
FT	DOMAIN	431	C345C (DOMAIN C).	581	C345C (DOMAIN C).
FT	SITE	507	CELL ATTACHMENT SITE	509	POTENTIAL.

FT	DISULFID	91	BY SIMILARITY.
FT	DISULFID	262	BY SIMILARITY.
FT	DISULFID	271	BY SIMILARITY.
FT	DISULFID	284	BY SIMILARITY.
FT	DISULFID	283	BY SIMILARITY.
FT	DISULFID	292	BY SIMILARITY.
FT	DISULFID	295	BY SIMILARITY.
FT	DISULFID	318	BY SIMILARITY.
FT	DISULFID	327	BY SIMILARITY.
FT	DISULFID	340	BY SIMILARITY.
FT	DISULFID	345	BY SIMILARITY.
FT	DISULFID	348	BY SIMILARITY.
FT	DISULFID	360	BY SIMILARITY.
FT	DISULFID	378	BY SIMILARITY.
FT	DISULFID	381	BY SIMILARITY.
FT	DISULFID	383	BY SIMILARITY.
FT	DISULFID	400	BY SIMILARITY.
FT	DISULFID	402	BY SIMILARITY.
FT	DISULFID	411	BY SIMILARITY.
FT	DISULFID	414	BY SIMILARITY.
FT	DISULFID	428	BY SIMILARITY.

FT	DISULFID	453	521	BY SIMILARITY.
FT	DISULFID	468	578	BY SIMILARITY.
FT	CARBOHYD	88	88	POTENTIAL.
FT	CARBOHYD	103	103	POTENTIAL.
FT	CARBOHYD	394	394	POTENTIAL.
FT	CARBOHYD	540	540	POTENTIAL.
SO	SEQUENCE	581 AA:	65106 MW: 5B6D2272	CRC32:

Query Match 7.1%; Score 148; DB 1; Length 581;
Best Local Similarity 27.0%; Pred. No. 1.06e-10;
Matches 30; Conservative 28; Mismatches 46; Indels 7; Gaps 6;

[illegible]

RESULT	3	STANDARD;	PRT;	606 AA.
ID	NETL_CHICK			
AC	Q90922;			
DT	01-NOV-1997 (REL. 35, CREATED)			
DT	01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)			
DT	01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)			
DE	NETRIN-1 PRECURSOR.			
OS	GALLUS GALLUS (CHICKEN).			
OC	EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; AVES; NEOGNATHAE.			
OC	GALLIFORMES.			

SEQUENCE FROM N.A.
RC STRAIN=WHITE LEGHORN; TISSUE=EMBRYONIC BRAIN;
RX MEDLINE; 94340732.
RA SERAFINI T., KENNEDY T.E., MIRZAYAN C., JESSELL T.M.,
RA TESSIER-LAVIGNE M.;
RL CELL 78-409-424 (1994).
CC -1- FUNCTION: NETRINS CONTROL GUIDANCE OF CNS COMMISSURAL AXONS AND
PERIPHERAL MOTOR AXONS.
CC -1- SUBCELLULAR LOCATION: EXTRACELLULAR.
CC -1- SIMILARITY: CONTAINS ONE LAMININ N-TERMINAL DOMAIN (DOMAIN VI).
CC -1- SIMILARITY: CONTAINS 3 LAMININ EGF-LIKE DOMAINS.
CC -1- SIMILARITY: CONTAINS ONE C345C DOMAIN.
CC -1- SIMILARITY: TO CARNOBARDITIS ELEGANS UNC-6 PROTEIN

DR EMBL: L34549; G529419; -.

DR PROSITE; PS0022; EGF_1; 2.

DR PROSITE; PS01248; LAMININ_TYPE_EGF; 3.

KW GLYCOPROTEIN; EXTRACELLULAR MATRIX; SIGNAL; LAMININ EGF-LIKE DOMAIN

NW	REFEAL.	1	25	POTENTIAL.
FT	SIGNAL	26	606	NETRIN-1.
FT	CHAIN	26	286	LAMININ N-TERMINAL (DOMAIN VI).
FT	DOMAIN	287	455	3 X LAMININ EGF-LIKE REPEATS (LAMININ DOMAIN V-LIKE).

FT	343	405	LAMININ EGF-LIKE 2.
DOMAIN			
FT	406	455	LAMININ EGF-LIKE 3.
DOMAIN			
FT	456	606	C345C (DOMAIN C).
DOMAIN			
FT	532	534	CELL ATTACHMENT SITE (POTENTIAL).
FT SITE			

FT	DISULFID	287	296	BY SIMILARITY.
FT	DISULFID	289	306	BY SIMILARITY.
FT	DISULFID	308	317	BY SIMILARITY.
FT	DISULFID	320	340	BY SIMILARITY.
FT	DISULFID	343	352	BY SIMILARITY.
FT	DISULFID	345	370	BY SIMILARITY.
FT	DISULFID	373	382	BY SIMILARITY.
FT	DISULFID	385	403	BY SIMILARITY.
FT	DISULFID	406	418	BY SIMILARITY.
FT	DISULFID	408	425	BY SIMILARITY.
FT	DISULFID	427	436	BY SIMILARITY.
FT	DISULFID	439	453	BY SIMILARITY.

[illegible]

```

RL MOL. CELL. BIOL. 9:3850-3859(1989).
CC -1- FUNCTION: PHOSPHORYLATES A WIDE RANGE OF SUBSTRATES INCLUDING
CC RIBOSOMAL PROTEIN S6.
CC -1- TISSUE SPECIFICITY: INTESTINE, THYMUS, AND LUNG.
CC -1- SIMILARITY: THE N-TERMINUS SHOWS HIGH SEQUENCE SIMILARITY TO THE
CC REGIONS OF PROTEIN KINASE C, THE CATALYTIC SUBUNIT OF CAMP-
CC DEPENDENT PROTEIN KINASE, AND CGMP-DEPENDENT PROTEIN KINASE. THE
CC REMINDER OF S6 KINASE II SHOWS HIGH SEQUENCE SIMILARITY TO THE
CC CATALYTIC SUBUNIT OF PHOSPHORYLASE B KINASE.
CC EMBL; M28489; G556322; -.
CC PIR; B32571; B32571.
CC HSSP; P05132; 1CTP.
CC MGD; MG1:104558; RPS6KAL.
CC PROSITE; PS00107; PROTEIN_KINASE_ATP; 2.
CC PROSITE; PS00108; PROTEIN_KINASE_ST; 2.
CC PROSITE; PS50011; PROTEIN_KINASE_DOM; 2.
CC TRANSFERASE; SERINE/THREONINE-PROTEIN KINASE; ATP-BINDING;
CC REPEAT; MULTIGENE FAMILY.
CC FT DOMAIN 62 310 PROTEIN KINASE 1.
CC FT NP_BIND 407 664 PROTEIN KINASE 2.
CC FT NP_BIND 68 76 ATP (BY SIMILARITY).
CC FT BINDING 94 94 ATP (BY SIMILARITY).
CC FT ACT_SITE 187 187 BY SIMILARITY.
CC FT NP_BIND 413 421 ATP (BY SIMILARITY).
CC FT BINDING 436 436 ATP (BY SIMILARITY).
CC FT ACT_SITE 524 524 BY SIMILARITY.
CC SEQUENCE 724 AA; 81594 MW; F961A261 CRC32;

Query Match 4.8%; Score 99; DB 1; Length 724;
Best Local Similarity 35.2%; Pred.No. 5.72e-02;
Matches 19; Conservative 13; Mismatches 18; Indels 4; Gaps 4;

Db 489 KI-LRQKFFSEREASVFLHTISKTVEYLSQGVVHRDLKPSNLTLYVDESGNPEC 541
QY 177 KIKVKEITYINRDTKILETKSTII-Y-KLVGVSEKDLKRS-VLWLDKSLQCTC 227

RESULT 9
ID VQ07_BPT4 STANDARD; PRT; 1032 AA.
ID C P19061;
DT 01-NOV-1990 (REL. 16, CREATED)
DT 01-NOV-1990 (REL. 16, LAST SEQUENCE UPDATE)
DT 01-NOV-1990 (REL. 16, LAST ANNOTATION UPDATE)
DE BASEPLATE STRUCTURAL PROTEIN GP7.
GN 7.
OS BACTERIOPHAGE T4.
OC VIRIDAE; DS-DNA NONENVELOPED VIRUSES; MYOVIRIDAE.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=D;
RX MEDLINE; 90384864.
RA EFIMOV V.P., PRILIPOV A.G., MESYANZHINOV V.V.;
RL NUCLEIC ACIDS RES. 18:5313-5313(1990).
CC -1- FUNCTION: STRUCTURAL COMPONENT OF THE BASEPLATE.
CC EMBL; X15907; G15323; -.
CC PIR; JQ0657; G7BPT4.
CC STRUCTURAL PROTEIN.
CC SEQUENCE 1032 AA; 119214 MW; 39F5B2D1 CRC32;

Query Match 4.7%; Score 98; DB 1; Length 1032;
Best Local Similarity 39.0%; Pred.No. 8.18e-02;
Matches 16; Conservative 8; Mismatches 15; Indels 2; Gaps 2;

Db 368 MDSICKVFAIIGVEYETLNAPRTSKII-DSADKGIYVLN 407
QY 166 METLCKNDFAIKI-KVKEITYINRDTKILETKSKTIYKLN 205

RESULT 10
ID ION3_CARAU STANDARD; PRT; 520 AA.
ID C P18520;
DT 01-NOV-1990 (REL. 16, CREATED)
DT 01-NOV-1990 (REL. 16, LAST SEQUENCE UPDATE)

```

[illegible]

CC STRANDED DNA IN AN ATP-DEPENDANT MANNER AND ALSO CATALYZES THE
 CC INTERCONVERSION OF OTHER TOPOLOGICAL ISOMERS OF DOUBLE-STRANDED
 CC DNA RINGS, INCLUDING CATENANES AND KNOTTED RINGS.
 CC -!- CATALYTIC ACTIVITY: ATP-DEPENDENT BREAKAGE, PASSAGE AND REJOINING
 CC OF DOUBLE-STRANDED DNA.
 CC -!- SUBUNIT: MADE UP OF TWO CHAINS. THE A CHAIN IS RESPONSIBLE FOR DNA
 CC BREAKAGE AND REJOINING; THE B CHAIN CATALYZES ATP HYDROLYSIS. THE
 CC ENZYME FORMS AN A2B2 TETRAMER.
 DR EMBL: D73436; G1322131; -.
 DR EMBL: D73421; G1322101; -.
 DR PROSITE: PS00177; TOPOISOMERASE_II; 1.
 KW TOPOISOMERASE: ISOMERASE; ATP-BINDING; ANTIBIOTIC RESISTANCE.
 FT NON_TER 1
 FT NON_TER 116 117
 FT NON_TER 216 216
 SQ SEQUENCE 216 AA; 23775 MW; E572BEE0 CRC32;
 Query Match 4.3%; Score 90; DB 1; Length 216;
 Best Local Similarity 26.3%; Pred. No. 1.29e+00;
 Matches 20; Conservative 23; Mismatches 26; Indels 7; Gaps 7;
 Db 79 LARRLESLNAGVRIVLDERVALEHFDLEVLSEKSL-DIAGLPGLA-DCOEKD 136
 QY 176 LKIKVKEITYINRDKIIL-ETK-S-KTIYKLN-GVSEDLKKSVLWLKDSLOCTCEEMN 231
 Db 137 PALSLEYLVEGDSAGG 152
 QY 232 -DINAPYLVMGOKGG 246
 RESULT 14
 ID VIRF_SHIDY STANDARD; PRT; 262 AA.
 AC Q04248;
 DT 01-JUN-1994 (REL. 29, CREATED)
 DT 01-JUN-1994 (REL. 29, LAST SEQUENCE UPDATE)
 DT 01-NOV-1995 (REL. 32, LAST ANNOTATION UPDATE)
 DE VIRULENCE REGULON TRANSCRIPTIONAL ACTIVATOR VIRF.
 GN VIRF.
 OS SHIGELLA DYSENTERIAE, SHIGELLA FLEXNERI, AND SHIGELLA SONNEI.
 OG PLASMID 210 KB INVASION.
 OC PROKARYOTA; GRACILICUTES; SCOTOBACTERIA; FACULTATIVELY ANAEROBIC RODS;
 OC ENTEROBACTERIACEAE.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC SPECIES=S.DYSENTERIAE; STRAIN=CG097;
 RA YAO R., REDDY L.V., PALCHAUDHURI S.;
 RL SUBMITTED (MAR-1991) TO EMBL/GENBANK/DBJ DATA BANKS.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC SPECIES=S.FLEXNERI; STRAIN=2A;
 RX MEDLINE; 87032409.
 RA SAKAI T., SASAKAWA C., MAKINO S., YOSHIKAWA M.;
 RL INFECT. IMMUN. 54:395-402(1986).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC SPECIES=S.SONNEI;
 RX MEDLINE; 89212881.
 RA KATO J.I., ITO K.I., NAKAMURA A., WATANABE H.;
 RL INFECT. IMMUN. 57:1391-1398(1989).
 RN [4]
 RP SIMILARITY TO ARAC FAMILY.
 RX MEDLINE; 92326642.
 RA DORMAN C.J.;
 RL MOL. MICROBIOL. 6:1575-1575(1992).
 CC -!- FUNCTION: TRANSCRIPTIONAL ACTIVATOR OF THE VIRB GENE WHICH IS
 CC ITSELF AN ACTIVATOR OF THE IPAABC VIRULENCE REGULON.
 CC -!- SIMILARITY: BELONGS TO THE ARAC/XLYS FAMILY OF TRANSCRIPTIONAL
 CC REGULATORS.
 DR EMBL: X58464; G46954; -.
 DR EMBL: M29172; G152801; -.
 DR EMBL: X16661; G47067; -.
 DR PIR: S14646; S14646.
 DR PIR: A47605; A47605.

DR PIR: A60105; A60105.
 DR PROSITE: PS00041; HTH_ARAC_FAMILY_1; 1.
 DR PROSITE: PS01124; HTH_ARAC_FAMILY_2; 1.
 KW PLASMID; VIRULENCE; TRANSCRIPTION REGULATION; ACTIVATOR; DNA-BINDING.
 FT DNA_BIND 177 196 H-T-H MOTIF (BY SIMILARITY).
 SQ SEQUENCE 262 AA; 30549 MW; 1139B38 CRC32;
 Query Match 4.3%; Score 89; DB 1; Length 262;
 Best Local Similarity 25.5%; Pred. No. 1.79e+00;
 Matches 13; Conservative 21; Mismatches 15; Indels 2; Gaps 2;
 Db 37 TLTIDEQIAFIERNICINYSIKKSDSINPFETISLDNRLLLSIIRIMEPI 87
 QY 175 ALKIKVKEITYINRDKIILETK-SKTIYKLVGS-ERDLKKSVLWLKDSL 223
 RESULT 15
 ID KSGA_MYCPN STANDARD; PRT; 263 AA.
 AC P75113;
 DT 01-NOV-1997 (REL. 35, CREATED)
 DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)
 DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
 DE DIMETHYLADENOSINE TRANSFERASE (EC 2.1.1.-) (S-ADENOSYLMETHIONINE-6-N',
 DE N'-ADENOSYL(RRNA) DIMETHYLTRANSFERASE) (16S RNA DIMETHYLASE) (HIGH
 DE LEVEL KASUGAMYCIN RESISTANCE PROTEIN KSGA) (KASUGAMYCIN
 DE DIMETHYLTRANSFERASE).
 GN KSGA.
 OS MYCOPLASMA PNEUMONIAE.
 OC PROKARYOTA; TENERICUTES; MOLLICUTES; MYCOPLASMA; MYCOPLASMATALES;
 OC MYCOPLASMATACEAE.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 29342 / M129;
 RX MEDLINE; 97105885.
 RA HIMMELREICH R., HILBERT H., PLAGENS H., PIRKL E., LI B.-C.,
 RA HERRMANN R.;
 RL NUCLEIC ACIDS RES. 24:4420-4449(1996).
 CC -!- FUNCTION: SPECIFICALLY DIMETHYLATES TWO ADJACENT ADENOSINES IN THE
 CC LOOP OF A CONSERVED HAIRPIN NEAR THE 3'-END OF 16S RNA IN THE 30S
 CC PARTICLE. ITS INACTIVATION LEADS TO KASUGAMYCIN RESISTANCE
 CC (BY SIMILARITY).
 CC -!- SIMILARITY: STRONG, TO OTHER BACTERIAL KSGA.
 CC -!- SIMILARITY: BELONGS TO THE RNA ADENINE N-6-METHYLTRANSFERASE
 CC FAMILY.
 DR EMBL: AE000017; G1673824; -.
 DR PROSITE: PS01131; RNA_A_DIMETH; 1.
 KW MRNA PROCESSING; TRANSFERASE; METHYLTRANSFERASE;
 KW ANTIBIOTIC RESISTANCE.
 SQ SEQUENCE 263 AA; 29794 MW; 357AA00D CRC32;
 Query Match 4.3%; Score 90; DB 1; Length 263;
 Best Local Similarity 37.0%; Pred. No. 1.29e+00;
 Matches 20; Conservative 11; Mismatches 20; Indels 3; Gaps 3;
 Db 52 KLQLPYHGIEL-DKRLAEVLLVNEILTEQLTIGDALKONLDQYFDDTPIPLCG 104
 QY 26 NLQL-CHGIEYQNMRLFN-LLGHETMKVELEQAGAWIPLVMKQCHPDTRKFLCS 77
 Search completed: Tue Aug 4 10:42:21 1998
 Job time : 25 secs.

M O S E R H
***** (TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

MPsrch_pp protein - protein database search, using Smith-Waterman algorithm
Run on: Tue Aug 4 10:42:39 1998; MasPar time 18.68 Seconds
Tabular output not generated. 619.822 Million cell updates/sec

Title: >US-08-848-439-3
Description: (1-275) from US08848439.pep
Perfect Score: 2073
Sequence: 1 SARGFLFGQPDFSYKRSNC.....WQKGQREFKRSIRKLOC 275

Scoring table: PAM 150
Gap 11
Searched: 140555 seqs, 42109429 residues
Post-processing: Minimum Match 0%
Listing first 45 summaries
Database: sptrembl6
1:sp_fungi 2:sp_human 3:sp_invertebrate 4:sp_mammal
5:sp_mhc 6:sp_organelle 7:sp_phase 8:sp_plant
9:sp_bacteria 10:sp_rodent 11:sp_virus 12:sp_vertebrate
13:sp_unclassified

Statistics: Mean 44.539; Variance 72.874; scale 0.611
Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES				Pred. No.	
Result No.	Score	Query Match Length	ID	Description	
1	2067	99.7	295 10	O08862	SECRETED FRIZZLED REL 0.00e+00
2	2065	99.6	295 10	O05297	SECRETED APOPTOSIS REL 0.00e+00
3	2062	99.5	295 10	P97299	STROMAL CELL DERIVED F 0.00e+00
4	1377	66.4	206 2	O14778	SECRETED APOPTOSIS REL 0.00e+00
5	913	44.0	314 2	O14779	SECRETED APOPTOSIS REL 8.59e-193
6	911	43.9	313 2	O00346	SECRETED FRIZZLED-RELA 2.74e-192
7	902	43.5	308 4	O19116	FRZA PRECURSOR 5.17e-190
8	901	43.5	317 2	O14780	SECRETED APOPTOSIS REL 9.26e-190
9	890	42.9	314 10	O08861	SECRETED FRIZZLED REL 5.58e-187
10	729	35.2	307 12	O42397	CRESCENT 1.63e-146
11	398	19.2	588 3	O16147	MOM-5 1.59e-65
12	396	19.1	572 10	O61090	FRIZZLED HOMOLOG 7 (TR 4.72e-65
13	394	19.0	591 2	O00144	FRIZZLED HOMOLOG 8 (TR 1.41e-64
14	394	19.0	685 10	O61091	FRIZZLED GENE PRODUCT 1.09e-62
15	386	18.6	565 2	O14332	FRIZZLED PROTEIN HOMOL 1.09e-62
16	386	18.6	570 10	O08464	FRIZZLED PROTEIN HOMOL 1.09e-62
17	386	18.6	641 10	O08463	FRIZZLED PROTEIN HOMOL 1.09e-62
18	374	18.0	585 2	O13467	TRANSMEMBRANE RECEPTOR 7.34e-60
19	369	17.8	537 10	O61088	FRIZZLED HOMOLOG 4 (TR 1.10e-58
20	357	17.2	525 3	O10662	FRIZZLED PROTEIN HOMOL 7.10e-56

21 355 17.1 694 3 Q94916 DF22. 2.08e-55
22 347 15.7 586 3 Q24760 TISSUE POLARITY PROTEI 1.53e-53
23 337 16.3 558 3 Q94132 TRANSMEMBRANE RECEPTOR 3.21e-51
24 317 15.3 666 10 Q61086 FRIZZLED-3. 1.33e-46
25 309 14.9 664 12 Q42579 FRIZZLED 3. PROTEIN. 9.05e-45
26 295 14.2 605 3 O16190 F27E11.3 PROTEIN. 1.41e-41
27 293 14.1 325 4 Q95117 FRZB PRECURSOR. 4.00e-41
28 291 14.0 325 2 O00181 FRITZ. 1.14e-40
29 289 13.9 325 2 Q99686 FRZB PRECURSOR. 3.22e-40
30 289 13.9 325 2 Q92765 FRIZZLED HOMOLOG 6 (TR 4.34e-39
31 284 13.7 709 10 Q61089 FRIZZLED RELATED PROTE 2.06e-38
32 281 13.6 348 10 Q35222 SECRETED FRIZZLED-RELA 4.59e-37
33 275 13.3 323 10 P97401 FRPHE. 7.70e-37
34 274 13.2 368 2 O14877 FRZB PRECURSOR. 2.83e-35
35 267 12.9 319 12 P79993 FRIZZLED. 7.90e-35
36 265 12.8 318 12 P79936 FRIZZLED. 2.59e-29
37 240 11.6 261 10 O08570 FRITZ (FRAGMENT). 1.04e-19
38 195 9.4 641 2 O00520 CARBOXYPEPTIDASE Z PRE 3.16e-16
39 178 8.6 505 10 Q60569 PROCOLLAGEN, TYPE XVII 3.16e-16
40 178 8.6 562 10 Q60672 PROCOLLAGEN, TYPE XVII 3.16e-16
41 178 8.6 1774 10 Q62001 ALPHA-1(XVII) COLLAGE 4.01e-06
42 125 6.0 602 12 O42203 NETRIN 1. 6.01e-06
43 124 6.0 794 12 O42224 SMOOTHENED (FRAGMENT). 1.34e-05
44 122 5.9 787 2 Q99835 SMOOTHENED. 1.34e-05
45 122 5.9 793 10 P97698 SMOOTHENED. 1.34e-05

ALIGNMENTS

RESULT 1
ID O08862 PRELIMINARY: PRT; 295 AA.
AC O08862;
DT 01-JUL-1997 (TREMBLREL. 04, CREATED)
DT 01-JUL-1997 (TREMBLREL. 04, LAST SEQUENCE UPDATE)
DT 01-JUL-1997 (TREMBLREL. 04, LAST ANNOTATION UPDATE)
DE SECRETED FRIZZLED RELATED PROTEIN SRP-2.
SF SRP2.
GN MUS MUSCULUS (MOUSE).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC EUKARYOTA; RODENTIA.
RN [1]
RP SEQUENCE FROM N.A.
RA RATTNER A., HSIEH J.C., SMALLWOOD P.M., GILBERT D.J.,
RA COPELAND N.G., JENKINS N.A., NATHANS J.;
RL PROC. NATL. ACAD. SCI. U.S.A. 94:2859-2863(1997).
DR EMBL: U88567; G1946343; .
SQ SEQUENCE 295 AA; 33469 MW; 236B0DDA CRC32;

Query Match 99.7%; Score 2067; DB 10; Length 295;
Best Local Similarity 99.3%; Pred. No. 0.00e+00;
Matches 273; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Db 21 SARGFLFGQPDFSYKRSNCPIPANLOLCHGIEYQNNRPNLLGHETMKEVLEQAGAWI 80
QY 1 SARGFLFGQPDFSYKRSNCPIPANLOLCHGIEYQNNRPNLLGHETMKEVLEQAGAWI 60
Db 81 PLVMKQCHPDTKKFLCSLFPVCLDDDETIOCHSLCVQKRCAPVMSAFGFPWDM 140
QY 61 PLVMKQCHPDTKKFLCSLFPVCLDDDETIOCHSLCVQKRCAPVMSAFGFPWDM 120
Db 141 ECDRFQNDLCIPLASSDHLPLATEAPKVEACKTNEDNDIMETLCKNDFALKIV 200
QY 121 ECDRFQNDLCIPLASSDHLPLATEAPKVEACKTNEDNDIMETLCKNDFALKIV 180
Db 201 KEITYINRDTKILLETSKTIYKLVGSEKLVKSVLWKDSLOCTCEEMNDINAPILVM 260
QY 181 KEITYINRDTKILLETSKTIYKLVGSEKLVKSVLWKDSLOCTCEEMNDINAPILVM 240
Db 261 GOKQGGELVITSVKRWQKQREFKRSIRKLOC 295
QY 241 GOKQGGELVITSVKRWQKQREFKRSIRKLOC 275

```

RESULT 2
ID O35297 PRELIMINARY; PRT; 295 AA.
AC O35297;
DT 01-JAN-1998 (TREMBLREL. 05, CREATED)
DT 01-JAN-1998 (TREMBLREL. 05, LAST SEQUENCE UPDATE)
DT 01-JAN-1998 (TREMBLREL. 05, LAST ANNOTATION UPDATE)
DE SECRETED APOPTOSIS RELATED PROTEIN 1.
GN SARPL
OS MUS MUSCULUS (MOUSE).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC EUTHERIA; RODENTIA.
RN [1]
RA MELKONYAN H., CHANG W.C., SHAPIRO J.P., MAHADEVAPPA M.,
RA FITZPATRICK P.A., KIEFER M.C., TOMEI D.L., UMANSKY S.R.;
RL SUBMITTED (SEP-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL; AF017989; G2415421;
SQ SEQUENCE 295 AA; 33483 MW; 0E4C8A8E CRC32;

Query Match 99.6%; Score 2065; DB 10; Length 295;
Best Local Similarity 98.9%; Pred. No. 0.00e+00;
Matches 272; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Db 21 SARGLFQGPDPFSYKRSNCKPIPANLQCHGIEYQNMRLPNLLGHETMKEVLEQAGAWI 80
QY 1 SARGLFQGPDPFSYKRSNCKPIPANLQCHGIEYQNMRLPNLLGHETMKEVLEQAGAWI 60

Db 81 PLVMKQCHPDTKKFLCSLFAPVCLDDLDDETIOQCHSLCVQVKDRCAPVMSAFGFPWPDML 140
QY 61 PLVMKQCHPDTKKFLCSLFAPVCLDDLDDETIOQCHSLCVQVKDRCAPVMSAFGFPWPDML 120

Db 141 ECDRFPQDNDLCIPLASSDHLHPATEAPKVCACKTKNEDDNDIMETLCKNDFALKIKV 200
QY 121 ECDRFPQDNDLCIPLASSDHLHPATEAPKVCACKTKNEDDNDIMETLCKNDFALKIKV 180

Db 201 KEITYINRDTKIILETKSKTIYKLVGVSERDLKSKVLWLDKSLQCTCEEMNDINAPYLVM 260
QY 181 KEITYINRDTKIILETKSKTIYKLVGVSERDLKSKVLWLDKSLQCTCEEMNDINAPYLVM 240

Db 261 GQGGELVITSVKRWQKQREFKRISRSIRKLOC 295
QY 241 GQGGELVITSVKRWQKQREFKRISRSIRKLOC 275

Db 81 PLVMKQCHPDTKKFLCSLFAPVCLDDLDDETIOQCHSLCVQVKDRCAPVMSAFGFPWPDML 140
QY 61 PLVMKQCHPDTKKFLCSLFAPVCLDDLDDETIOQCHSLCVQVKDRCAPVMSAFGFPWPDML 120

Db 141 ECDRFPQDNDLCIPLASSDHLHPATEAPKVCACKTKNEDDNDIMETLCKNDFALKIKV 200
QY 121 ECDRFPQDNDLCIPLASSDHLHPATEAPKVCACKTKNEDDNDIMETLCKNDFALKIKV 180

Db 201 KEITYINRDTKIILETKSKTIYKLVGVSERDLKSKVLWLDKSLQCTCEEMNDINAPYLVM 260
QY 181 KEITYINRDTKIILETKSKTIYKLVGVSERDLKSKVLWLDKSLQCTCEEMNDINAPYLVM 240

Db 261 GQGGELVITSVKRWQKQREFKRISRSIRKLOC 295
QY 241 GQGGELVITSVKRWQKQREFKRISRSIRKLOC 275

RESULT 3
ID P97299 PRELIMINARY; PRT; 295 AA.
AC P97299;
DT 01-MAY-1997 (TREMBLREL. 03, CREATED)
DT 01-MAY-1997 (TREMBLREL. 03, LAST SEQUENCE UPDATE)
DT 01-JAN-1998 (TREMBLREL. 05, LAST ANNOTATION UPDATE)
DE STROMAL CELL DERIVED FACTOR 5 (SDF5).
GN SDF5.
OS MUS MUSCULUS (MOUSE).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC EUTHERIA; RODENTIA.
RN [1]
RA SHIROZU M., TADA H., TASHIRO K., NAKAMURA T., LOPEZ N.D.,
RA NAZAREA M., HAMADA T., SATO T., NAKANO T., HONJO T.;
RL GENOMICS 37:273-280(1996).
DR EMBL; D50462; G1747302;
DR MGD; MGI:108078; SDF5.
SQ SEQUENCE 295 AA; 33501 MW; 329888C8 CRC32;

Query Match 99.5%; Score 2062; DB 10; Length 295;
Best Local Similarity 98.9%; Pred. No. 0.00e+00;
Matches 272; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Db 21 SARGLFQGPDPFSYKRSNCKPIPANLQCHGIEYQNMRLPNLLGHETMKEVLEQAGAWI 80
QY 1 SARGLFQGPDPFSYKRSNCKPIPANLQCHGIEYQNMRLPNLLGHETMKEVLEQAGAWI 60

Db 81 PLVMKQCHPDTKKFLCSLFAPVCLDDLDDETIOQCHSLCVQVKDRCAPVMSAFGFPWPDML 140

```

```

QY 61 PLVMKQCHPDTKKFLCSLFAPVCLDDLDDETIOQCHSLCVQVKDRCAPVMSAFGFPWPDML 120
Db 141 ECDRFPQDNDLCIPLASSDHLHPATEAPKVCACKTKNEDDNDIMETLCKNDFALKIKV 200
QY 121 ECDRFPQDNDLCIPLASSDHLHPATEAPKVCACKTKNEDDNDIMETLCKNDFALKIKV 180
Db 201 KEITYINRDTKIILETKSKTIYKLVGVSERDLKSKVLWLDKSLQCTCEEMNDINAPYLVM 260
QY 181 KEITYINRDTKIILETKSKTIYKLVGVSERDLKSKVLWLDKSLQCTCEEMNDINAPYLVM 240
Db 261 GQGGELVITSVKRWQKQREFKRISRSIRKLOC 295
QY 241 GQGGELVITSVKRWQKQREFKRISRSIRKLOC 275

RESULT 4
ID O14778 PRELIMINARY; PRT; 206 AA.
AC O14778;
DT 01-JAN-1998 (TREMBLREL. 05, CREATED)
DT 01-JAN-1998 (TREMBLREL. 05, LAST SEQUENCE UPDATE)
DT 01-JAN-1998 (TREMBLREL. 05, LAST ANNOTATION UPDATE)
DE SECRETED APOPTOSIS RELATED PROTEIN 1 (FRAGMENT).
GN SARPL.
OS HOMO SAPIENS (HUMAN).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC EUTHERIA; PRIMATES.
RN [1]
RA MELKONYAN H., CHANG W.C., SHAPIRO J.P., MAHADEVAPPA M., FITZPATRICK P.A.,
RA KIEFER M.C., TOMEI D.L., UMANSKY S.R.;
RL SUBMITTED (SEP-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL; AF017986; G2415415;
FT NON_TER 206
SQ SEQUENCE 206 AA; 23215 MW; F465B68B CRC32;

Query Match 66.4%; Score 1377; DB 2; Length 206;
Best Local Similarity 98.4%; Pred. No. 0.00e+00;
Matches 185; Conservative 0; Mismatches 1; Indels 2; Gaps 1;

Db 21 SARGLFQGPDPFSYKRSNCKPIPANLQCHGIEYQNMRLPNLLGHETMKEVLEQAGAWI 80
QY 1 SARGLFQGPDPFSYKRSNCKPIPANLQCHGIEYQNMRLPNLLGHETMKEVLEQAGAWI 60

Db 81 PLVMKQCHPDTKKFLCSLFAPVCLDDLDDETIOQCHSLCVQVKDRCAPVMSAF--PWPDM 138
QY 61 PLVMKQCHPDTKKFLCSLFAPVCLDDLDDETIOQCHSLCVQVKDRCAPVMSAFGFPWPDML 120

Db 139 ECDRFPQDNDLCIPLASSDHLHPATEAPKVCACKTKNEDDNDIMETLCKNDFALKIKV 198
QY 121 ECDRFPQDNDLCIPLASSDHLHPATEAPKVCACKTKNEDDNDIMETLCKNDFALKIKV 180

Db 199 KEITYINR 206
QY 181 KEITYINR 188

RESULT 5
ID O14779 PRELIMINARY; PRT; 314 AA.
AC O14779;
DT 01-JAN-1998 (TREMBLREL. 05, CREATED)
DT 01-JAN-1998 (TREMBLREL. 05, LAST SEQUENCE UPDATE)
DT 01-JAN-1998 (TREMBLREL. 05, LAST ANNOTATION UPDATE)
DE SECRETED APOPTOSIS RELATED PROTEIN 2.
GN SARP2.
OS HOMO SAPIENS (HUMAN).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC EUTHERIA; PRIMATES.
RN [1]
RA MELKONYAN H., PROCHAZKA V., CHANG W.C., UMANSKY S.R.;
RL SUBMITTED (SEP-1997) TO EMBL/GENBANK/DBJ DATA BANKS.

```



```
DR EMBL: AF017987; G2415417; -.
SQ SEQUENCE 314 AA; 35411 MW; 61E261B5 CRC32;

Query Match 44.0%; Score 913; DB 2; Length 314;
Best Local Similarity 42.6%; Pred. No. 8,55e-193;
Matches 112; Conservative 67; Mismatches 77; Indels 7; Gaps 5;

Db 51 FYTKPPQCVDPADLRLCHNVGKKMVLPLNLEHETMAEVKQOASSWVPLLNKNCHAGTQ 110
QY 13 FSKRSNCKPIPANLQCHGIEYQNMRLPNLLGHETMKVELEQAGAWIPLVNMQCHPDTK 72

Db 111 VFLCSLFAPVCLD---RPIYPCRWLCEAVRDSCEPVMOFFGFYWPMLKCDKFP-EGDVC 166
QY 73 KFLCSLFAPVCLDLDLDTETQPCSLCVQKRCAPVMSAFGFPWPDMLCEDRFPQDNLDLC 132

Db 167 IAMTPNPTEASKPOGTTVCPDCNDELKSEA-IIIEHLCASEFALRMKIKVEKKENGDKKI 225
QY 133 IPLASSDHLLPATEAPKVCCEACKNKNDNDIMETLCKNDFAIKIKVEITYINRDTKI 192

Db 226 V-PKKKKPL-KLGPIKKDLKLVLYLKNAGDCPCHQDNLNLSHFLIMGRKVKSOYLLTA 283
QY 193 ILETSKTIYKLVNGVSRDLKKSVLWLDLSQCTCEEMNDINAPYLVGMQKOGGELVITS 252

Db 284 IHKWDKKNKEFKNMKKNHEC 306
QY 253 VVRWQKGOREFKRISIRKLQOC 275

RESULT 6
ID O00546 PRELIMINARY; PRT; 313 AA.
AC O00546;
DT 01-JUL-1997 (TREMBLREL. 04, CREATED)
DT 01-JUL-1997 (TREMBLREL. 04, LAST SEQUENCE UPDATE)
DE SECRETED FRIZZLED-RELATED PROTEIN.
OS HOMO SAPIENS (HUMAN).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC EUTHERIA; PRIMATES.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=LUNG;
RA FINCH P.W., HE X., KELLEY M.J., UREN A., SCHAUDIES R.P.,
RA POPESCU N.C., RUDIKOFF S., AARONSON S.A., VARMS H.E., RUBIN J.S.;
RL SUBMITTED (JUN-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL: AF001900; G2213819; -.
SQ SEQUENCE 313 AA; 35314 MW; A3CD04FF CRC32;

Query Match 43.9%; Score 911; DB 2; Length 313;
Best Local Similarity 42.6%; Pred. No. 2,74e-192;
Matches 112; Conservative 67; Mismatches 77; Indels 7; Gaps 5;

Db 50 FYTKPPQCVDPADLRLCHNVGKKMVLPLNLEHETMAEVKQOASSWVPLLNKNCHAGTQ 109
QY 13 FSKRSNCKPIPANLQCHGIEYQNMRLPNLLGHETMKVELEQAGAWIPLVNMQCHPDTK 72

Db 110 VFLCSLFAPVCLD---RPIYPCRWLCEAVRDSCEPVMOFFGFYWPMLKCDKFP-EGDVC 165
QY 73 KFLCSLFAPVCLDLDLDTETQPCSLCVQKRCAPVMSAFGFPWPDMLCEDRFPQDNLDLC 132

Db 166 IAMTPNPTEASKPOGTTVCPDCNDELKSEA-IIIEHLCASEFALRMKIKVEKKENGDKKI 224
QY 133 IPLASSDHLLPATEAPKVCCEACKNKNDNDIMETLCKNDFAIKIKVEITYINRDTKI 192

Db 225 V-PKKKKPL-KLGPIKKDLKLVLYLKNAGDCPCHQDNLNLSHFLIMGRKVKSOYLLTA 282
QY 193 ILETSKTIYKLVNGVSRDLKKSVLWLDLSQCTCEEMNDINAPYLVGMQKOGGELVITS 252

Db 283 IHKWDKKNKEFKNMKKNHEC 305
QY 253 VVRWQKGOREFKRISIRKLQOC 275

RESULT 7
```

```
ID O19116 PRELIMINARY; PRT; 308 AA.
AC O19116;
DT 01-JAN-1998 (TREMBLREL. 05, CREATED)
DT 01-JAN-1998 (TREMBLREL. 05, LAST SEQUENCE UPDATE)
DT 01-JAN-1998 (TREMBLREL. 05, LAST ANNOTATION UPDATE)
DE FRZA PRECURSOR.
OS BOS TAURUS (BOVINE).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC EUTHERIA; ARTIODACTYLA.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=HEART;
RA DUPLAA C., D'AMORE P.A.;
RL SUBMITTED (AUG-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL: U85945; G2337937; -.
KW SIGNAL.
FT SIGNAL.
SQ SEQUENCE 308 AA; 34763 MW; 9B74A92F CRC32;

Query Match 43.5%; Score 902; DB 4; Length 308;
Best Local Similarity 42.2%; Pred. No. 5,17e-190;
Matches 111; Conservative 66; Mismatches 79; Indels 7; Gaps 5;

Db 45 FYTKPPQCVDPADLRLCHNVGKKMVLPLNLEHETMAEVKQOASSWVPLLNKNCHIGTQ 104
QY 13 FSKRSNCKPIPANLQCHGIEYQNMRLPNLLGHETMKVELEQAGAWIPLVNMQCHPDTK 72

Db 105 VFLCSLFAPVCLD---RPIYPCRWLCEAVRDSCEPVMOFFGFYWPMLKCDKFP-EGDVC 160
QY 73 KFLCSLFAPVCLDLDLDTETQPCSLCVQKRCAPVMSAFGFPWPDMLCEDRFPQDNLDLC 132

Db 161 IAMTPNPTEASKPOGTTVCPDCNDELKSEA-IIIEHLCASEFALRMKIKVEKKENGDKKI 219
QY 133 IPLASSDHLLPATEAPKVCCEACKNKNDNDIMETLCKNDFAIKIKVEITYINRDTKI 192

Db 220 V-PKKKKPL-KLGPIKKDLKLVLYLKNAGDCPCHQDNLNLSHFLIMGRKVKSOYLLTA 277
QY 193 ILETSKTIYKLVNGVSRDLKKSVLWLDLSQCTCEEMNDINAPYLVGMQKOGGELVITS 252

Db 278 IHKWDKKNKEFKNMKKNHEC 300
QY 253 VVRWQKGOREFKRISIRKLQOC 275

RESULT 8
ID O14780 PRELIMINARY; PRT; 317 AA.
AC O14780;
DT 01-JAN-1998 (TREMBLREL. 05, CREATED)
DT 01-JAN-1998 (TREMBLREL. 05, LAST SEQUENCE UPDATE)
DT 01-JAN-1998 (TREMBLREL. 05, LAST ANNOTATION UPDATE)
DE SECRETED APOPTOSIS RELATED PROTEIN 3.
GN SARP3.
OS HOMO SAPIENS (HUMAN).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC EUTHERIA; PRIMATES.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=PANCREAS;
RA MELKONIAN H., CHANG W.C., SHAPIRO J.P., MAHADEVAPPA M., FITZPATRICK P.A.,
RA KIEFER M.C., TOMBI D.L., UNANSKY S.R.;
RL SUBMITTED (SEP-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL: AF017988; G2415419; -.
SQ SEQUENCE 317 AA; 35577 MW; 2CD69373 CRC32;

Query Match 43.5%; Score 901; DB 2; Length 317;
Best Local Similarity 45.2%; Pred. No. 9,26e-190;
Matches 114; Conservative 58; Mismatches 74; Indels 6; Gaps 4;

Db 46 SYSKPPQCLDIPADLPLCHTGVYKRMVLPLNLEHESLAEVKQOASSWVPLLNKLRCHSDTO 105
QY 14 SY-KRSNCKPIPANLQCHGIEYQNMRLPNLLGHETMKVELEQAGAWIPLVNMQCHPDTK 72

Db 106 VFLCSLFAPVCLD---RPIYPCRWLCEAVRDSCEPVMOFFGFYWPMLKCDKFP-EGDVC 162
```


Qy 142 LPA 144

Search completed: Tue Aug. 4 10:44:04 1998

Job time : 85 secs.

THIS PAGE BLANK (user)

M O S E R E H

(TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (C) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

MPsrch_nn n.a. - n.a. database search, using Smith-Waterman algorithm

Run on: Wed Aug 5 03:42:10 1998; MasPar time 1087.57 Seconds
Tabular output not generated.

Title: >rev-trans-seq-2
Description: (1-885) from translate.seq
Perfect Score: 734
N.A. Sequence: 1 ATGYNCAAGGCGGCCGWS.....SNATMNGNAARYTNCARTGY 885
Comp: TACRANGTYCCGGGCCGNCWS.....SNTANKCNTTYRANGTYACR

Scoring table: TABLE default
Gap 6

Nmatch STD : Dbase 0; Query 0

Searched: 532261 seqs, 918536377 bases x 2

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: emb155

Database: 1:em_ba 2:em_htg 3:em_hum1 4:em_hum2 5:em_in 6:em_om
7:em_or 8:em_ov 9:em_pat 10:em_pl 11:em_ro
12:gb_ba 13:gb_htg 14:gb_in 15:gb_om 16:gb_ov 17:gb_pat
18:gb_ph 19:gb_pl 20:gb_pri 21:gb_pr2 22:gb_ro 23:gb_st
24:gb_sts 25:gb_sy 26:gb_un 27:gb_v1

Statistics: Mean 14.843; Variance 13.039; scale 1.138

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	722	98.4	2028	22	MMU8567 Mus musculus secreted	0.00e+00
2	722	98.4	2031	22	AF017989 Mus musculus secreted	0.00e+00
3	720	98.1	1776	22	D50462 Mouse SDF5 mRNA, compl	0.00e+00
4	477	65.0	882	21	AF017986 Homo sapiens secreted	3.23e-200
5	144	19.6	1984	21	AF017988 Homo sapiens secreted	7.81e-44
6	134	18.3	2094	21	AF017987 Homo sapiens secreted	2.20e-39
7	131	17.8	2075	21	AF001900 Homo sapiens secreted	4.69e-38
8	131	17.8	4659	21	AF056087 Homo sapiens secreted	4.69e-38
9	123	16.8	2659	22	MMU8566 Mus musculus secreted	1.56e-34
10	122	16.6	4240	15	BTU85945 Bos taurus frizzled-re	4.29e-34
11	122	16.6	7218	17	I66494 Sequence 14 from paten	4.29e-34
12	99	13.5	2075	16	AF006508 Gallus gallus crescent	3.74e-24
13	89	12.1	2421	21	MMU43321 Mus musculus putative	6.16e-20
14	75	10.2	2334	20	HSU43318 Human putative transme	3.50e-14
15	69	9.4	1881	22	AF054623 Mus musculus frizzled-	8.77e-12

16	64	8.7	1770	14	CEU43316	Caenorhabditis elegans	8.02e-10
17	63	8.6	1181	16	AF059570	Xenopus laevis secretate	1.96e-09
18	63	8.6	2831	22	MMU43317	Mus musculus putative	1.96e-09
19	62	8.4	2184	21	HSU82169	Human frizzled homolog	4.76e-09
20	61	8.3	1727	16	AF031831	Gallus gallus 7-transm	1.15e-08
21	61	8.3	2085	14	DMU65589	Drosophila melanogaste	1.15e-08
22	61	8.3	2328	16	AF031830	Gallus gallus 7-transm	1.15e-08
23	59	8.0	2260	22	MMU43320	Mus musculus putative	6.67e-08
24	56	7.6	1913	14	CEU63557	Caenorhabditis elegans	8.99e-07
25	56	7.6	4540	22	RATFR22H	Rattus norvegicus Dros	8.99e-07
26	54	7.4	1923	21	HUMFRIZ	Human frizzled gene pr	4.97e-06
27	53	7.2	215	17	I28278	Sequence 5 from paten	1.16e-05
28	52	7.1	1845	14	AF013953	Caenorhabditis elegans	2.69e-05
29	52	7.1	1912	22	RATFR2H	Rattus norvegicus Dros	2.69e-05
30	52	7.1	10772	14	AF012089	Drosophila melanogaste	2.69e-05
31	52	7.1	34576	14	CET23D8	Caenorhabditis elegans	2.69e-05
32	50	6.8	10772	14	AF012089	Drosophila melanogaste	1.43e-04
33	49	6.7	565	17	E04076	gDNA encoding envelope	3.25e-04
34	47	6.4	215	17	I28278	Sequence 5 from paten	1.66e-03
35	46	6.3	2624	22	MMU43205	Mus musculus frizzled-	3.72e-03
36	46	6.3	7218	17	I66494	Sequence 14 from paten	3.72e-03
37	45	6.1	25700	14	CELF27E11	Caenorhabditis elegans	8.25e-03
38	44	6.0	354	14	OFU89259	Oxytricha fallax 57kd	1.82e-02
39	43	5.9	354	14	OFU89259	Oxytricha fallax 57kd	3.98e-02
40	43	5.9	565	17	E04076	gDNA encoding envelope	3.98e-02
41	43	5.9	16741	14	PFSC03095	Plasmodium falciparum	3.98e-02
42	43	5.9	95908	14	PFSC03010	Plasmodium falciparum	3.98e-02
43	40	5.4	370	14	OFU89262	Oxytricha fallax 57kd	3.97e-01
44	40	5.4	2478	9	AF63310	Sequence 3 from paten	3.97e-01
45	40	5.4	3342	21	AB012911	Homo sapiens mRNA for	3.97e-01

ALIGNMENTS

RESULT	1	MMU8567	2028 bp	mRNA	22-APR-1997
LOCUS		Mus musculus secreted	frizzled related protein	sFRP-2 (Sfrp2)	mRNA.
DEFINITION		complete cds.			
ACCESSION		U88567			
NID		91946342			
KEYWORDS		house mouse.			
SOURCE		Mus musculus			
ORGANISM		Eukaryotae; Mitochondrial eukaryotes; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
REFERENCE		1 (bases 1 to 2028)			
AUTHORS		Rattner, A., Hsieh, J.-C., Smallwood, P.M., Gilbert, D.J., Copeland, N.G., Jenkins, N.A. and Nathans, J.			
TITLE		A family of secreted proteins contains homology to the cysteine-rich ligand-binding domain of frizzled receptors			
JOURNAL		Proc. Natl. Acad. Sci. U.S.A. 94 (7), 2859-2863 (1997)			
MEDLINE		97250455			
REFERENCE		2 (bases 1 to 2028)			
AUTHORS		Rattner, A., Hsieh, J.-C., Smallwood, P.M., Gilbert, D.J., Copeland, N.G., Jenkins, N.A. and Nathans, J.			
TITLE		Direct Submission			
JOURNAL		Submitted (05-FEB-1997) Molecular Biology and Genetics, HHMI/Johns Hopkins University School of Medicine, 725 North Wolfe Street, PC1B 805, Baltimore, MD 21205			
FEATURES		Location/Qualifiers			
Source		1. .2028			
		/organism="Mus musculus"			
		/db_xref="taxon:10090"			
		/chromosome="3"			
		/map="near Fgg"			
		236. .1123			
gene		/gene="sfrp2"			
		236. .1123			
CDS		/gene="Sfrp2"			
		/codon_start=1			
		/product="secreted frizzled related protein sFRP-2"			
		/db_xref="PID:g1946343"			


```
Db 614 GTGAAGACCGCTGGCCCGCGTGTGTCGCTTCCGCTTCCCTGCGCAGACATGCTG 673
Qy 361 GTNARGATGNTGTCNCCTGATGWSGCTTGGTTCCTTGGCCGATGATGYN 420
Db 674 GAGTCGACCGTTTCCCGAGGACACAGACCTCTGCTATCCCTCGCTAGTAGCAGCAC 733
Qy 421 GARTGYGATGNTTCCNCARGAAYGAYTNTGYATNCCNTNGCNWSNNGAYCAY 480
Db 734 CTCCTGCGCGCCACAGAGAAGCTCCCAAGGTGTGTGAAGCTGCAAAACCAAGATGAG 793
Qy 481 YTNYNCCGNCACNGARGCNCCTNAARGTNTGYGARGCTNTGYAARAAYAAAYGAY 540
Db 794 GACGACACGACATCATGATAAACCCTTTCTGAAAATGACTTCGCACTGAAAATCAAAGTG 853
Qy 541 GAYGAYAAATGATNATGARGACNTTNTGYAARAAYGAYTTCGCTNATNATNARGTN 600
Db 854 AAGGAGATTAACGTATACATCAACAGACACCAAGATCATCTCGAGACAAGAGACACC 913
Qy 601 AARGARATNACNTATATNAAYMGNGAYACNAARATNATNTNGARACNAARNAARCN 660
Db 914 ATTACACCTGAACGGCTGTCGGAAGGACCTGAAGAAATCGTGTGGCTCAAA 973
Qy 661 ATNTAAYATTAAYGNGTNGSNGARMGNGAYTNTAARAARWSNGTNTTGGYTNAAR 720
Db 974 GACAGCTGCACTGACCTGTGAGGAGATGAACGACATCAACGCTCCGCTATCTGCTCATG 1033
Qy 721 GAYWSYNTNARTGYACNTGYGARGARATGAAYGATNAAYGNCNTAYTNTNGTATG 780
Db 1034 GGACAGAACGAGCGGCGGAACTGTTGATCACTCGCTGAAACGGTGGCAGAGGCCAG 1093
Qy 781 GGNCAARACRGGNGGNGARNTNGTNTNACNWSNGTNAARMGNTGGCARAARGNCAR 840
Db 1094 AGAGAGTTCAAGCGATCTCCGCGACCATCCGCAAGCTCAATGC 1138
Qy 841 MNGARTTYAARMGNATNWSNMGNSNATNMGNAARYTNCARTGY 885

RESULT
LOCUS D50462 1776 bp mRNA ROD 24-DEC-1996
DEFINITION Mouse SDF5 mRNA, complete cds.
ACCESSION D50462
NID g1747301
KEYWORDS SDF5.
SOURCE Mus musculus cell_line:ST-2 cDNA to mRNA, clone_lib:phage (lgt22a)
library, ST-2 directional 1.
ORGANISM Mus musculus
Eukaryotae; Mitochondrial eukaryotes; Metazoa; Chordata;
Vertebrata; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae;
Murinae; Mus.
REFERENCE
1 (sites)
Shirozu,M., Tada,H., Tashiro,K., Nakamura,T., Lopez,N.D.,
Nazarea,M., Hamada,T., Sato,T., Nakano,T. and Honjo,T.
Characterization of novel secreted and membrane proteins isolated
by the signal sequence trap method
Genomics 37 (3), 273-280 (1996)
97092876
REFERENCE
2 (bases 1 to 1776)
Shirozu,M., Tada,H., Nakamura,T., Nelson,L.D., Martina,N.,
Hamada,T., Sato,T., Tashiro,K., Nakano,T. and Honjo,T.
Isolation of novel genes encoding for secreted or membrane proteins
using signal sequence trap
unpublished (1995)
3 (bases 1 to 1776)
Shirozu,M.
Direct Submission
Submitted (01-MAY-1995) to the DDBJ/EMBL/GenBank databases. Michio
Shirozu, Kyoto University, Faculty of Medicine, Department of
Medical Chemistry, Yoshida, Sakyo-ku, Kyoto, Kyoto 606, Japan
(E-mail:kondo@viru1.virus.kyoto-u.ac.jp, Tel:81-75-753-4377,
Fax:81-75-753-4388)
LOCATION/Qualifiers
1. 1776
/organism="Mus musculus"
```

```

/db_xref="taxon:10090"
/cell_line="ST-2"
/clone_lib="phage (lgt22a) library, ST-2 directional 1"
14..901
/gene="SDF5"
14..901
/gene="SDF5"
/codon_start=1
/db_xref="PID:d1009688"
/db_xref="PID:g1747302"
/translation="MPGPAASLLLLVLAHSHCCGASGLFLFGQPDFSYKRSNCKPFP
ANQLCHGIEYQNMRLNPLGHETWKEVLEOAGAWIPLVMKQHPDKKELCSLFAPV
LLDDEETIQPHSLQKQVDRCAPVMSAFGFPWMDLECDREDFQDDLCIPLASSDH
LLDPATEAPKVCCEACKTKNEDNDIMETLCKNDFALKIKVEITFIYINRDKIILETKS
KTIYKLVGVSERDLKSVLKLKLSLOCTCEEMNDINAPYLVMGQKGGLVITSVKRW
OKGOREFKRISRSIRKLQC"
BASE COUNT 433 a 483 c 405 g 455 t
ORIGIN
Query Match 98.1%; Score 720; DB 22; Length 1776;
Best Local Similarity 58.1%; Pred. No. 0.00e+00;
Matches 514; Conservative 213; Mismatches 158; Indels 0; Gaps 0;
Db 14 ATGCCGCGGGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 73
Qy 1 ATGYTNCARGGNCNGNWSNNTYNTNTNTNTNTNTNTNTNTNTNTNTNTNTNTNTNTNT 60
Db 74 TCGGCGCGTGGCTTCTCTTCTGCGCCAGCCGACTTCTCTTCTGCGCCAGCCGACTTC 133
Qy 61 WSNCGNMNGGNTTNTTNTTNTTNTTNTTNTTNTTNTTNTTNTTNTTNTTNTTNTTNT 120
Db 134 AAGCCCATCCCGCCCAACCTGCGCTGTGCCAGCGATCGAGTACCAAGCAATCGCGCTG 193
Qy 121 AARCCNATNCCNCAAYTNCARTNTGYCAYGNATNGARTAYCARAAYATGNGNTN 180
Db 194 CCCAACCTGCTGGGCCACGACCATGAAGGAGTGTGGAGCAGCGCGCGCTGGATT 253
Qy 181 CCNAAAYTNTYNGGNCAYGARACNATGAARGNTNYTNGARCGNCGNGCNTGGATN 240
Db 254 CGGCTGTCATGAACGAGTGCACCCGCGCAGACACCAAGATTCCTGCTGCTGCTCC 313
Qy 241 CCNTYNGTATGAARCACTGTCAYCCNGAYACNAARAATTTTNTGYWSNTNTTYG 300
Db 314 CCGTCTGCTGCTGCGACGCTAGATGAGACCATCCAGCGCTGCTGCTGCTGCTGCTG 373
Qy 301 CCNGTNTGYTNGAYGAYTNGAYGARACNATNCARCCNTGYCAYWSNTNTGYCTN 360
Db 374 GTGAAGACCGCTGCGCCCGCTGATGTCGCTTCCGCTTCCCTGCGCAGACATGCTG 433
Qy 361 GTNAARGAYMGNTGYCNCNGTNTATGWSNCGNTTGYGNTTTCNTGCGCNGAYATG 420
Db 434 GAGTCGACCGTTTCCCGCAGGACACACCATCTGCACTCCCTCCCTAGTAGCAGCAC 493
Qy 421 GARTGYGATGNTTCCNCARGAAYGAYTNTGYATNCCNTNGCNWSNNGAYCAY 480
Db 494 CTCCTGCGCGCCACAGAGAAGCTCCCAAGGTGTGTGAAGCTGCAAAACCAAGATGAG 553
Qy 481 YTNYNCCGNCACNGARGCNCCTNAARGTNTGYGARGCTNTGYAARAAYAAAYGAY 540
Db 554 GACGACACGACATCATGAAAACCTTTGTAAAATGACTTCGCACTGAAAATCAAAGTG 613
Qy 541 GAYGAYAAATGATNATGARGACNTTNTGYAARAAYGAYTTCGCTNATNATNARGTN 600
Db 614 AAGGAGATTAACGTATACATCAACAGACACCAAGATCATCTGAGACAAGAGACACC 673
Qy 601 AARGARATNACNTATATNAAYMGNGAYACNAARATNATNTNGARACNAARNAARCN 660
Db 674 ATTTCAAGCTGAACGGCTGTCCGAAAGGACCTGAAGAAATCCGCTGCTGCTGCTCAA 733
Qy 661 ATNTAAYATTAAYGNGTNGSNGARMGNGAYTNTAARAARWSNGTNTTGGYTNAAR 720
Db 734 GACAGCTGCACTGCGACCTGTGAGGAGATGAACGACATCAACGCTCCGCTATCTG 793
Qy 721 GAYWSYNTNARTGYACNTGYGARGARATGAAYGATNAAYGNCNTAYTNTNGTATG 780
```

BASE COUNT
ORIGIN

Query Match		19.6%	Score 144;	DB 21;	Length 1984;
Best Local Similarity		39.4%	Pred. No. 7.81e-44;		
Matches		298;	Conservative 131;	Mismatches 313;	Indels 15; Gaps 10;
Db	350	CTCCTACTCAAGCGCGCAGTGGCTTGACATCCTCGCCACCTCGCGCTCTGCCACAC	409		
Qy	96	YTTYNSNTAYAAAGMWSNAATYGAARCCNATNCCNCAAYTYNCARTYNTGYCAYGG	155		
Db	410	GGTGGCTACAAAGCGCATCGGCTGCCAACCTGCTGGAGCAGAGCCTGGCCGAAGT	469		
Qy	156	NATNGARTAYCARAYATGCMGNTNCCNAAYYTYNTGNCNCAARGACNATGAARGART	215		
Db	470	GAAGCAGCAGGCGACGCTGCTGCCCTGCTGCTGCGCCAAAGCGCTGCCACTCGGATACGA	529		
Qy	216	NYTNGARCARGCNGGNCNGTGGATNCCNTNGTNGATGAARCAARTGYCAYCCNGAYACNA	275		
Db	530	GGTCTTCCTGCTCGCTCTTTGGCGCCGCTGTGCT- --CGACC- GG- C- ---CCATCTA	580		
Qy	276	RAARTTYTNGYWSNTTYTTCNCCNCTGTYTNGAYGAYTYNGAYGARACNATNCA	335		
Db	581	CCCGTGGCCTCGCTGTGGAGCGCGTGGCGCGGCTGCGCGCGCTCATGGAGGCCTA	640		
Qy	336	RCNTGYCAYWSNTYGTNCARGTNAARGAYWNTGYGNCNCTGATGWSNCGNTT	395		
Db	641	CGGCTTCCCTGGCTGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	700		
Qy	396	YGGNTTYCCTGGCNGAYATGTYNGARTGYGAYWNTTYCCNCAARGAYAYGAYTYNTG	455		
Db	701	CATCGCCGT- GC- AGTTCGGC- ACCTGCCCGCCACCCTGCTGCT- --CGACC- GG- C- ---CATCTA	757		
Qy	456	YATNCCNTYTNCGNWSNWSNGAYCAYTYTNTCCNCCNACNGARGARCCNCAARGTNTG	515		
Db	758	CCGCCAGTGTGAGATGGAGCAC- A- G- TCGTACGGCGCTCATGGAGCAGATGTGCTCCAG	814		
Qy	516	YGARGCNTGYAARAAYARAAYGAYGAYAAAYATNATGGARACNTYNTGYAARA	575		
Db	815	TGACTTTTGTGCAAAATGCGCATCAAGGAGATCAAGATAGAGAAATGGGACCGGAAGCT	874		
Qy	576	YGAYTTYGNTYNAARATNAARGTNAARGATNACNTATYATNAAAYMNGAYACNAARAT	635		
Db	875	GATTGGAGCCCAAGAAAGAGAGCTGCTCAAGCGCGGCCCTGAGCGCAGGACAC	934		
Qy	636	NATYNTNGARACNAARWSNAARACNATNTAYAAATNAYTNGAYGNGTNGNARNGMNGAYYT	695		
Db	935	CAAGCGCTGTGCTGCACATGAAGATGGCGCGGCTGCGCGCTGCCACACTGGACAG	994		
Qy	696	NAARARWSNGTYNTGGYTNAAARGAYWSYNTNARTGYACNTGYGARARATGAAYGA	755		
Db	995	CCTGGCGGCGAGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1054		
Qy	756	YATNAYGCNCCNTAYTYNTGATGGNCARAARCAARGGNGGNGARYTNGTATNACNWS	815		
Db	1055	CCTCACCCTGGGACAGAAGAATAAGGATGAAG	1091		
Qy	816	NGTNAARMGNTGGCARAARGGNCARMGNGARTTYAAR	852		
RESULT 6					
LOCUS	AF017987	2094 bp	mrna	PRI	21-SEP-1997
DEFINITION	Homo sapiens secreted apoptosis related protein 2 (SARP2) mRNA, complete cds.				
ACCESSION	AF017987				
KEYWORDS	NID 92415416				
SOURCE	human.				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryotae; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
REFERENCE	1 (bases 1 to 2094)				
AUTHORS	Melkonyan,H., Prochazka,V., Chang,W.C. and Umansky,S.R.				
TITLE	SARPs - a new family of proteins that regulate apoptosis				
JOURNAL	Unpublished				
REFERENCE	2 (bases 1 to 2094)				

AUTHORS	Melkonyan,H., Prochazka,V. and Umansky,S.R.				
TITLE	Direct Submission				
JOURNAL	Submitted (11-AUG-1997) LXR Biotechnology, Inc., 1401 Marina Way South, Richmond, CA 94804, USA				
FEATURES	Location/Qualifiers				
source	1..2094				
gene	/organism="Homo sapiens"				
	/db_xref="taxon:9606"				
	/chromosome="8"				
	/tissue_type="heart"				
	1..2094				
CDS	/gene="SARP2"				
	/note="similar to frizzled-like receptor, extracellular protein"				
	/codon_start=1				
	/product="secreted apoptosis related protein 2"				
	/db_xref="PID:g2415417"				
	/translation="MGIRSEGGRRGAALGVLLAALAVGSASEYDVVSQSDIG PQSGREYTPPQCVDIPADRLCHNVYKRYKVLPLELHEHETNAEYKQASSWVPLLN KNCAGTQVFLCSLFPVCDLDRPIPCRWLCEAVRDSCEPMOFFGYWPEMLCKDKF PEGDVCIAMTPPNPTEASKPQGTTCPPCDNELKSEAIIEHLCASFEALRMKTKVEVK ENGDKKIVPKKPKLPIKKDLKLVLYLKNAGDCPCPDLDNLSHFLIMGRKVK				
BASE COUNT	484 a	598 c	631 g	381 t	
ORIGIN					
Query Match	18.3%	Score 134;	DB 21;	Length 2094;	
Best Local Similarity	39.5%	Pred. No. 2.20e-39;			
Matches	312;	Conservative 140;	Mismatches 317;	Indels 21;	Gaps 13;
Db	452	CTTCTACACCAAGCCACCTCAGTCGTGGCAGTCCCGCGGACCTGGCGGTGTCGCCACAA	511		
Qy	96	YTTYNSNTAYAAAGMWSNAATYGAARCCNATNCCNCAAYTYNCARTYNTGYCAYGG	155		
Db	512	CGTGGCTACAGAAGATGGTGTGCCCAACCTGCTGGAGCAGCAGACCATGGCGGAGGT	571		
Qy	156	NATNGARTAYCARAAATATGMGNTNCCNAAYYTYNTGNCAYGARACNATGAARGART	215		
Db	572	GAAGCAGCAGCCACGACTGGTGGCCCTGCTCAACAAGAACTGCCACGCGGCACCCA	631		
Qy	216	NYTNGARCAAGCNGGNGCNTGGATNCCNTNGTNGTAAARCAARTGYCAYCCNGAYACNA	275		
Db	632	GGTCTTCTCTGCTCGCTTTCGGCGCGCTGCTGCT- G- -GACC- GG- C- ---CCATCTA	682		
Qy	276	RAARTTYTNGYWSNYTNTYGCNCCNGTNGTYTNGAYGAYTYNGAYGARACNATNCA	335		
Db	683	CCGCTGTGCTGGCTCTCGAGCGCGCTGCGCGACTCGTGGAGCGGCTCATGCAAGTCTT	742		
Qy	336	RCNTGYCAYWSNYTNGTYGTCARGTNAARGAYMNGTYGCNCCNGTATGWSNGCNTT	395		
Db	743	CGGCTTCTACTGGCCGAGATGCTTAAGTGTGACAAAGTTTCCCCAGGGGGACG- TCT- -G	799		
Qy	396	YGGNTTYCCTGGCNGAYATGYTNGARTGYAYMNGNTTYCCNCARGAYAAAYGAYTYNTG	455		
Db	800	CATCGCCATGACGCGCCCAATCCACCGAGCCTCCAAGCCCAAGCAACCGGTGTG	859		
Qy	456	YATNCCNTYTNCGNWSNWSNGAYCAYTYTNTCCNCCNACNGARGARCCNCAARGTNTG	515		
Db	860	TCCTCCCTGTGACACAGAGT- TGAA- ATC- TGAGGCCCATCATTAACATCTCTGTGCCAG	916		
Qy	516	YGARGCNTGYAARAAYARAAYGAYGAYATNATGARGACNTNTGYAARA	575		
Db	917	CGAGTTTGCAGTGGATGAATAAAGAAGTGAATAAAGAAAATGGCGNACAGAGAT	976		
Qy	576	YGAYTTYCNTNAAATNAARGTNAARGARATNACNTATYATNAAAYMNGNAYACNAAAT	635		
Db	977	TGTCGCCAAGA- -AGAAG- AAGCCCTG- -AAGTGGGGCCCATCAAGAAGAGGACCT	1030		
Qy	636	NATYNTNGARCAARWSNAARACNATNTAYAAATNAYGNGTNGNWSNGARMNGAYYT	695		
Db	1031	GAAGAAGCTTGTGCTGTACCTGAAGAATGGGGCTGACTGTCCCTGCCACCACTGGACAA	1090		

AUTHORS		Melkonyan,H., Prochazka,V. and Umansky,S.R.	
TITLE		Direct Submission	
JOURNAL		Submitted (11-AUG-1997) LXR Biotechnology, Inc., 1401 Marina Way South, Richmond, CA 94804, USA	
FEATURES		Location/Qualifiers	
source		1..2094	
		/organism="Homo sapiens"	
		/db_xref="taxon:9606"	
		/chromosome="8"	
		/tissue_type="heart"	
gene		1..2094	
		/gene="SARP2"	
CDS		303..1247	
		/gene="SARP2"	
		/note="similar to frizzled-like receptor, extracellular protein"	
		/codon_start=1	
		/product="secreted apoptosis related protein 2"	
		/translation="MGIRSEGGRRGAALGVLLAALLAVGSAEYDYVFSQSDIG PYQGRFYTKPPQCVDPADLRCHNHYKRWLPNLEHETMAEVQQAQSSWVPLLN KNCAGTQVFLCSLFAVPCVLDRIYPCRWLCEAVRDSCEPVWFOFFGYWPEMLKCDKF PEGDVCIAMTPNPPTASKPGQTTVCPDNLKSEALIEHLCASEFALRMKIEVKV ENGDKYVPPKKPLKPKKKLKLVLKNGADCPCHOLDNLSHFLIMGRKVK SQLLTATAIHKWDKKNEKFMKKMKHNECPTFQSVFK"	
BASE COUNT		484 a 598 c 631 g 381 t	
ORIGIN			
Query Match		18.3%; Score 134; DB 21; Length 2094;	
Best Local Similarity		39.5%; Pred. No. 2.20e-39;	
Matches		312; Conservative 140; Mismatches 317; Indels 21; Gaps 13;	
Db	452	CTTCTACACCAAGCCACCTCAGTGGTGGACATCCCCGGGACCTCGCGCTGTGCCACAA	511
Qy	96	YTYNSNTAYAAAGMWSNAATYGAARCCNATNCCNCAAYTYNCARTYNTGYCAYGG	155
Db	512	COTGGGCTACAGAAGATGGTGTGCCCAACCTGCTGGAGCAGCAGACCATGGCGAGGT	571
Qy	156	NATNGARTAYCARAAYATGCMGNTNCCNAAITYNTNGNCAYGARACNATGAARGGT	215
Db	572	GAAGCAGCAGGCCAGCAGCTGGTGGCTGCTCAACAAGAACTGCCAGCGCGGACCCA	631
Qy	216	NYTNGARCARGCNGCNGCTGATNCCNTNGTNGATGAARCAARTGYCAYCCNGAYACNA	275
Db	632	GGTCTTCTGCTGCTGCTTCTCGCGCCCTCTGCGCT-G-GACC-GG-C---CCATCTA	682
Qy	276	RAARTTYTNTGYWSNTTYTTCNCCNCTNTGYTNGAYGAYTYNGAYGARACNATNCA	335
Db	683	CCGCTGTGCTGCTGCTGCGAGCGCGTGGCGACTGCTGCGAGCGCGGTGTCATGCAGTCT	742
Qy	336	RCCNTGYCAYWSNTYNTGYTNCARGTNAARGAYMGTNTGYGNCNCGTNGATGWSNCGNT	395
Db	743	CGGCTTCTACTGGCCGAGATGCTTAAGTGTGACAAGTTTCCCGAGGGGACG-TCT--G	799
Qy	396	YGGNTTYCCTGGCCNGAYATGYTNGARTGYGAYMGTNTTYCCNCAARGAAYAYTYNTG	455
Db	800	CATGCCCATGACGCGCCCAATCCACCGACCTCCCAAGCCCTCCCAAGCCCAACGGGTG	859
Qy	456	YATNCCNTYTNCGNWSNWSNGAYCAYTYTNTCCNCCNACNGARGARCCNCAARGTNTG	515
Db	860	TCCTCCCTGTGACACAGAGT-TGAA-ATC-TGAGGCCATCATTTGAACATCTCTGTGCCAG	916
Qy	516	YGARGCNTGYAARAAYAAARAYGAYGAYATNATGGARACNATNTGYAARA	575
Db	917	CGAGTTTGCAGTGAAGTGAATAAACAAGTGAATAAAGAAATGCGGACAGAAAT	976
Qy	576	YGAYTYGCNTYNAARATNAARGTNAARGATNACNTATYATNAAITYMNGNGAYACNAARAT	635
Db	977	TGTCCCCAAGA--AGAAG-AAAGCCCTG--AAGTTGGGCCCCATCAAGAAGAGGACCT	1030
Qy	636	NATYTNNGARACNARWSNARACNATNTAYAAATNAYTNGGNGTNGNARNGNGAYYT	695
Db	1031	GAAGAAGCTGTGCTGTACCTGAAGAATGGGCTGACTGTCCCTGCCACAGCTGGACAA	1090

```

Qy 696 NAARAARWSNGTNTGGYTNAAAGAYWSNYTNCARTGYACNTGYGARGARATGAAYGA 755
Db 1091 CCTCAGCCACCACITCTCATCATCGGCCCGCAAGGTGAAGAGCCAGTCTGCTGACGGC 1150
Qy 756 YATNAAYGCNCNTAYTYTNGTATGGGNCARAARCARCGGNGGNGARYTNGTATNACNWS 815
Db 1151 CATCCACAAGTGAGACAAAACAAGAGGTTCACAAACATTCATGAAGAAAATGAAGAAA 1210
Qy 816 NGTNAARMGNTGGCARAARGNCARMGNARTTYAARMGNATNWSNMGNWSNATNMGNAA 875
Db 1211 CCATGAGTGC 1220
Qy 876 RYTNCARTGY 885

RESULT 7
LOCUS AF001900 2075 bp mRNA PRI 25-JUN-1997
DEFINITION Homo sapiens secreted frizzled-related protein mRNA, complete cds.
ACCESSION AF001900
NID 92213818
KEYWORDS human.
ORGANISM Homo sapiens
Eukaryotes; mitochondrial eukaryotes; Metazoa; Chordata;
Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Homiidae;
Homo.
REFERENCE 1 (bases 1 to 2075)
Finch, P.W., He X., Kelley, M.J., Uren, A., Schaudies, R.P.,
Popescu, N.C., Rudikoff, S., Aaronson, S.A., Varmus, H.E. and
Rubin, J.S.
Purification and Molecular Cloning of a Secreted, Frizzled-Related
Antagonist of Wnt Action
Proc. Natl. Acad. Sci. U.S.A. (1997) In press
2 (bases 1 to 2075)
Finch, P.W., Aaronson, S.A. and Rubin, J.S.
Direct Submission
Submitted (30-APR-1997) LCMB, DBS/NCI, 9000 Rockville Pike,
Bethesda, MD 20892, USA
3 (bases 1 to 2075)
Finch, P.W., Aaronson, S.A. and Rubin, J.S.
Direct Submission
Submitted (09-JUN-1997) LCMB, DBS/NCI, 9000 Rockville Pike,
Bethesda, MD 20892, USA
REMARK Sequence update by submitter
FEATURES
Location/Qualifiers
1..2075
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="8"
/map="8p11.1-p12"
/clone="HS1"
/dev_stage="embryonic"
/tissue_type="lung"
/cell_line="M426"
/cell_type="fibroblast"
303..1244
/notes="FRP"
/codon_start=1
/product="secreted frizzled-related protein"
/db_xref="PID:g2213819"
/translation="MGIGSEGGRRGALGVILGAALLAVGASAEYDVVSTQSDIGP
YQSGRYTPPCQVDIPADRLCHNVYGRKMYLPLEHETMAEYKQASVWPLLNK
NCHAGTQVLCFLDRPIYPCRWLCEAVRDSCEPMQFFGFYWPMLKCDKFP
EGDVCITAPPNATEASKPGQTVPCPCONELKSEAIIEHLGASFEALRMKIEVKKE
NGDKITVPKKPLKPIKPKDKDLKLVLYLKGADPCQDLNLSHFLINGRVKS
QYLLTAIHWDKKNKEFKMKNMKNHECTFQSVFK"
BASE COUNT 473 a 596 c 626 g 380 t
ORIGIN

Query Match 17.8%; Score 131; DB 21; Length 2075;
Best Local Similarity 39.4%; Pred. No. 4.69e-38;
Matches 311; Conservative 139; Mismatches 319; Indels 21; Gaps 14;

```

```

Db 449 CTTCTACCAACCCACCTCAGTGCCTGACATCCCGCGGACCTCGGCTGTGCCACAA 508
Qy 96 YTTWSTAYTAARMGWSNAAYTGYAARCCNATNCNCNAAIYTNCAINTYTCAYGG 155
Db 509 CGTGGCTACAAGAAGATGGTGTGCTGCCCAACCTGCTGGAGCAGGACCATCGCGAGGT 568
Qy 156 NATNGARTAYCARAAATATGMNTNCCNAAIYTNNGNCAIYGARACNATGAARGAT 215
Db 569 GAAGCAGCAGGCGAGCAGTGGTGGTCCCTGCTCAACAAGAATGCCACCCCGGACCCA 628
Qy 216 NYTNGARCARCGNGGNCNTGGATNCCNTNGTATGAACARTGYCAYCNGAYACNAA 275
Db 629 GGTCTTCCTCTGCTGCTCTTCGCGCCGCTGCTCCCT-G--GAAC-GG-C---CCATCTA 679
Qy 276 RAARTTYTNTGYWSNYTNTTYGCNCNGTNTGYTNGAYGAYTNGAYGARACNATNCA 335
Db 680 CCGGTGTGCTGCTGCTGCGAGCGCTGCGGACTCGTGGAGCCGGTCAATCGAGTCTT 739
Qy 336 RCCNTGYCAYWSNYTNTGYTNCARGTNAARGAYMNTGYCNCNGTNGATWSNGCNT 395
Db 740 CGGCTTCTACTGCGCGAGATGCTTAAGTGTGACAAGTTCCCGGAGGGGACG-TCTGCA 798
Qy 396 YGNTTYCCNTGCGCNGAYATGTYNGARTGYGAYMNTTYCCNARGAYAYGAYTNTG 455
Db 799 TCGCCATGAGCGCGCCCAATGCCACCGAA--GCCTCCAAGCCCAAGGCAC-AAACGGTGTG 856
Qy 456 YATNCCNYTNGCWNWSNNGAYCAYTYTNTCCNGCNCNARGARGCNCNARGTNTG 515
Db 857 TCCTCCTGTGACAACGAGT-TGAA-ATC-TGAGGCCATCATTTGAACATCTCTGTGCCAG 913
Qy 516 YGAGCNTGYAARAAYAAARAAYGAYGAYAYATNATGGARACNYTNTGYAARA 575
Db 914 CGAGTTTGCCTGAGCATGAARATAAAGACGTGAAAAGAAATGGCCACAGAGAT 973
Qy 576 YGATTYGNTTNAARTNAARTNAARGARATNACNTATATNAAYMNGNAYACNARAT 635
Db 974 TGTCCTCCAAAGA--AGAAG-AAGCCCTG---AAGTTGGGCGCCATCAAGAAGAAGACCT 1027
Qy 636 NATYNTGARACNAARWSNAARACNATNTAYAARYTNAAYGGNGTWSNGARMNGAYYT 695
Db 1028 GAAGAAGCTGTGCTGTACTGAAAGATGGGCTGACTGCTCCCTGCCACGAGTGGACAA 1087
Qy 696 NAARAARWSNGTYTNTGGYTNAAAGAYWSNYTNCARTGYACNTGYGARGARATGAAYGA 755
Db 1088 CPTCAGCCACCACTTCTCATCATGCGCGCAGAGGTGAAGAGCCTACTTGTGACGGC 1147
Qy 756 YATNAAYGCNCNTAYTYTNGTNGTNGGNCARAARCARCGGNGGNGARYTNGTATNACNWS 815
Db 1148 CATCCACAAGTGGGACAAAACAAGAGGTTCAAAACTTCATGAAGAAAATGAAGAAA 1207
Qy 816 NGTNAARMGNTGGCARAARGNCARMGNCARTTYAARMGNATNWSNMGNWSNATNMGNAA 875
Db 1208 CCATGAGTGC 1217
Qy 876 RYTNCARTGY 885

RESULT 8
LOCUS AF056087 4469 bp mRNA PRI 08-APR-1998
DEFINITION Homo sapiens secreted frizzled related protein mRNA, complete cds.
ACCESSION AF056087
NID g3033550
KEYWORDS human.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 4469)
Zhou, Z. and Wang, J.
Upregulation of human secreted Frizzled homologue in apoptosis and
its down regulation in breast tumors
JOURNAL Unpublished

```

```
REFERENCE 2 (bases 1 to 4469)
AUTHORS Zhou,Z. and Wang,J.
TITLE Direct Submission
JOURNAL Submitted (29-MAR-1998) Radiohemmet Research Laboratory, CKK,
Karolinska Institute, Solnavagen, Stockholm S-171 76, Sweden
FEATURES
source
1..4469
/organism="Homo sapiens"
/db_xref="taxon:9606"
303..1244
/function="putative Wnt antagonist"
/note="FRP; frizzled homolog; cell cycle and apoptosis
related; down regulated in breast tumorigenesis"
/codon_start=1
/product="secreted frizzled related protein"
/db_xref="PID:g303351"
/translation="MGIRSGRRGALGVLLAALLAVGSASEYDVYFQSDIGP
YOSGRFYTPQCVDPADLRICHNVGYKKMVLPLNLEHETMAEVQQAASWVPLLN
NCHAGTOVFLCSLFAVDLRPIYPCRWLCEAVRDSCEPVMOFFGYFPEMLKCDKFP
EGDVCIAMTPNATEASKPGQTVPCPCDNLKSEALIEHLCASEFALRMKIEVKKE
NGDKKIVPKKKPLKGLPKKKDLKLVLYLKGADCPCHOLDNLSHFLIMGRKVK
QYLLTAIHWDKKKEFNFKMKNHCEPTFQSVFK"
BASE COUNT 1121 a 1097 c 1120 g 1131 t
ORIGIN
Query Match 17.8%; Score 131; DB 21; Length 4469;
Best Local Similarity 39.4%; Pred.No. 4.69e-38;
Matches 311; Conservative 139; Mismatches 319; Indels 21; Gaps 14;
Db 449 CTTCTACACCAAGCCACCTCAGTGGTGGACATCCCGCGGACCTGCGGCTGTGCCACAA 508
QY :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||:
96 YTTYSNTAYAAARMGNWSNAATGYAARCCNATNCCGNCNAAYTNCARTYNTGYCAYGG 155
Db 509 CGTGGCTACAGAAGATGGTCTGCCACCACTGCTGGAGCAGACGACCATGGCGAGGT 568
QY :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||:
156 NATNGARTAYCARAAYATGMGNTYNTCCNAAYTNTNGNCNAYGARACNATGAARGGT 215
Db 569 GAAGCAGCAGGCGAGCAGTGGTGGTCCCTGCTCAACAAGAACTGCCAGCGCGGACCCA 628
QY :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||:
216 NYTNGARCAGCGNGCNGTGGATNCCNTNGTNGTNGATGAARCARTGYCAVCCNGAYACNA 275
Db 629 GGTCTTCCTCTGCTCGCTTTCGCGCCGCTGCTGCTT-G--GACC-GG-C-----CCATCTA 679
QY :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||:
276 RAARTTYTNTGYWSNYTNTTYGCNCCNGTNTGYTNGAYGAYTNGAYGARACNATNCA 335
Db 680 CCGGTGCTGCTGCTGCGAGCGCGTGGCGACCTGCTGGAGCGCGGTGCATCGAGTCTT 739
QY :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||:
336 RCNTGYCAYWSNYTNTGYTNCARTGAAGTNAARGAYMGNTGYGCNCCNGTNGTNGCNGT 395
Db 740 CGGCTTCTACTGGCGGAGATGTTAAGTGTGACAAAGTTCCTCCGGAGGGGACG-TCTGCA 798
QY :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||:
396 YGNTTYCCTGCGCNGAYATGTYNGARTGYGAYMGNTTYCCNCGARCAAYAYGAYTNTG 455
Db 799 TCGCCATGACGCGCCCAATGCCACCGAA-GCCTCCAAGCCCCCAAGGCAC-AACGGTGTG 856
QY :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||:
456 YATNCCNTNGCNSWSNGAYCAVYNTNTCCNGCNACNGARGARGCNCNAAAGTNTG 515
Db 857 TCCTCCTGTGACAGGAGT-TGAA-ATC-TGAGGCCATCATGAACATCTCTGTGCCAG 913
QY :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||:
516 YGARGCNTGYAARAAYAAAYGAYGAYAYAYATNATNGARACNYTNTGYAARA 575
Db 914 CGAGTTGACATGAGATGAATAAAGAACTGAATAAGAAATGGCGCAACAAGAT 973
QY :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||:
576 YGAYTYGCTYNAARATNAAGTNAARGARATNACNTAYATNAYMGNGATACNAARAT 635
Db 974 TGTCCCAAGA--AGAA--AAGCCCTG---AAGTTGGGCGCCATCAAGAAGGAGCT 1027
QY :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||:
636 NATNTNGARACNAARWSNAARACNATNTAYARYTNAAYGNGINGNWSNGARMNGAYT 695
Db 1028 GAAGAAGCTTGTGCTACCTGAAGAATGGGCGCTGACTGCTCCCTGCCACCACTGGACAA 1087
QY :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||:
696 NAARAARWSNGTNTYNTGGYTNAAARGAYWSNTNCAARTYACNTGYGARGARATGA 755
```

```
Db 1088 COTCAGCCACCACCTTCCTCATCATGGCCGCAAGAGTGAAGAGCCAGTACTTCTGACGCG 1147
QY :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||:
756 YATNAAYCNCNTAYTYNGTNGTGGNCARAAACARGCNGGNGARYTNGTNGTNGTNGTNG 815
Db 1148 CATCCACAGTGGGACAGAAAACAGAGTTCAAAACACTTCATGAAGAAAATGAAAAA 1207
QY :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||:
816 NCTNAARMGNTGGCAARAGGNCARMGNGARTTYAARMGNATNWSNMGNSNATNGNAA 875
Db 1208 CCATGAGTGC 1217
QY :|||:
876 RYTNCARTGY 885
RESULT 9
LOCUS MMU88566 2659 bp mRNA ROD 22-APR-1997
DEFINITION Mus musculus secreted frizzled related protein sFRP-1 (Sfrp1) mRNA,
complete cds.
ACCESSION U88566
NID 91946340
KEYWORDS house mouse.
SOURCE Mus musculus
ORGANISM
Eukaryota; mitochondria eukaryotes; Metazoa; Chordata;
Vertebrata; Mammalia; Euthera; Rodentia; Sciurognathi; Muridae;
Murinae; Mus.
REFERENCE 1 (bases 1 to 2659)
AUTHORS Rattner,A., Hsieh,J.-C., Smallwood,P.M., Gilbert,D.J.,
Copeland,N.G., Jenkins,N.A. and Nathans,J.
TITLE A family of secreted proteins contains homology to the
cysteine-rich ligand-binding domain of frizzled receptors
Proc. Natl. Acad. Sci. U.S.A. 94 (7), 2859-2863 (1997)
97250455
REFERENCE 2 (bases 1 to 2659)
AUTHORS Rattner,A., Hsieh,J.-C., Smallwood,P.M., Gilbert,D.J.,
Copeland,N.G., Jenkins,N.A. and Nathans,J.
TITLE Direct Submission
JOURNAL Submitted (05-FEB-1997) Molecular Biology and Genetics, HHMI/Johns
Hopkins University School of Medicine, 725 North Wolfe Street, PCTB
805, Baltimore, MD 21205
FEATURES
source
1..2659
/organism="Mus musculus"
/db_xref="taxon:10090"
/chromosome="8"
/map="between Plat and Egr1"
257..1201
/gene="Sfrp1"
257..1201
/gene="Sfrp1"
/codon_start=1
/product="secreted frizzled related protein sFRP-1"
/db_xref="PID:g1946341"
/translation="MGVRSARGGGAAGVLLAALLAALGAASAEYDVYFQSDIG
SYSGREYTKPQCVDPADLRICHNVGYKKMVLPLNLEHETMAEVQQAASWVPLLN
NCHMGTOVFLCSLFAVDLRPIYPCRWLCEAVRDSCEPVMOFFGYFPEMLKCDKFX
PEGDVCIAMTPNATEASKPGQTVPCPCDNLKSEALIEHLCASEFALRMKIEVKKE
ENGDKKIVPKKKPLKGLPKKKDLKLVLYLKGADCPCHOLDNLSHFLIMGRKVK
QYLLTAIHWDKKKEFNFKMKNHCEPTFQSVFK"
BASE COUNT 645 a 738 c 719 g 553 t 4 others
ORIGIN
Query Match 16.8%; Score 123; DB 22; Length 2659;
Best Local Similarity 38.9%; Pred.No. 1.56e-34;
Matches 307; Conservative 140; Mismatches 322; Indels 21; Gaps 14;
Db 406 CTTCTACCAAGCCCGCGAGTGGTGGACATCCCGTGGACCTGAGCTGTGCCACAA 465
QY :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||:
96 YTTYSNTAYAAARMGNWSNAATGYAARCCNATNCCNCAAYTNCARTYNTGYCAYGG 155
Db 466 CGTGGCTACAGAAGATGGTGTGCGCCCAACCTGCTGGAGCAGCAGACCATGGCAGAGGT 525
QY :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||:
156 NATNGARTAYCARAAYATGMGNTNCCNAAYTNTNGNCNAYGARACNATGAARGGT 215
```


RESULT	12	AF006508	2075 bp	mRNA	VRT
LOCUS		Gallus gallus	crescent	complete cds.	
DEFINITION		AF006508			
ACCESSION		92226371			
NID					
KEYWORDS					
SOURCE		chicken.			
ORGANISM		Gallus gallus			

RESULT	13				
LOCUS	MMU43321	2421 bp	DNA	ROD	24-FEB-1996
DEFINITION	Mus musculus putative transmembrane receptor (frizzled 8) gene, complete cds.				
ACCESSION	U43321				
NID	g1151259				
KEYWORDS	.				
SOURCE	house mouse.				
ORGANISM	Mus musculus				

Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata; Vertebrata; Eutheria; Rodentia; Sciurognathi; Myomorpha; Muridae; Murinae; Mus.

1 (bases 1 to 2421)

Wang,Y., Macke,J.P., Abella,B.S., Andreasson,K., Worley,P., Gilbert,D.J., Copeland,N.G., Jenkins,N.A. and Nathans,J.

A large family of putative transmembrane receptors homologous to the product of the Drosophila tissue polarity gene frizzled

J. Biol. Chem. 271 (8), 4468-4476 (1996)

96224032

2 (bases 1 to 2421)

Abella,B., Wang,Y., Macke,J.P. and Nathans,J.

Direct Submission

Submitted (14-DEC-1995) Jeremy Nathans, Molecular Biology and Genetics, Johns Hopkins Medical School, 725 N. Wolfe Street, Baltimore, MD 21205

Location/Qualifiers

1. .2421

/organism="Mus musculus"

/db_xref="taxon:10090"

/chromosome="18"

/map="between Tpl2 and Cdh2"

188. .2245

/gene="frizzled 8"

188. .2245

/gene="frizzled 8"

/note="putative transmembrane receptor"

/codon_start=1

/product="transmembrane receptor"

/db_xref="PID:g1151260"

/translation="MEWGYLLEVTSLAALAVLQRSSGAASAKELACQETIVPLCK GIGNYVYVNFQNDQDEAGLEWQFVPLVEIQSPDLKFFLCMTPIQLEDYKK PLPCRSVCERAKAGCAPLMROYGFAPWDMFCDRLPQGNPDTCMDYNRTDLTAA PEPRLPPLPPPGQSGSHSRPPCARPPHRCSSRGSGDAAPSPSGKARPP GGAAFCPCGCRAPMWSVSSEHPLNRYGQIQAACALPCHNPFSSQDRATVFWGLMSVLCFVFTATVFLDMEFYPERPIIFLSACYLFSVGYLVLRVAGHEK VACSGAGGAGGAGGAGAAAGAGAGASPGAGEYELGAVEQHRVETTPGA LCTVFLVLYVFGMASSIMWVLSLTWFLAAGKMGNAIAGYSOYFHLAAGVLSVK STAVLALSSVDGDPVAGICVYCNOSLDNLRGVFLAPLVLYLFGTFMFLAGVSLFRI RSVIQQGGPTTHKLEKIMRLGLFTVLYTPAAVVVACLVFEGHNPWEATHNCP CLRDLPQDARRPDPAVFMKLYFMCLVIGTSGVWWSGKTLSEWALCTRCCKWASKG AAVGAGAGSGSGGPGGPGGGGGGGGSLYSDVSTGLTWRSGTASSVSPKQMP L SQV"

BASE COUNT 403 a 772 c 779 g 467 t

ORIGIN

Query Match 12.1%; Score 89; DB 22; Length 2421;

Best Local Similarity 41.1%; Pred. No. 6.16e-20;

Matches 124; Conservative 47; Mismatches 131; Indels 0; Gaps 0;

Db 306 TGCCGTTGTGCAAGGCATCGGTTACAACTACACTTACATGCCCAACCAGTTCAACACG 365

Qy 140 TNCARYTNTGTCAYGNATGARTAYCARAAYATGMGNTNCCNAAYYNTYNGGNCAYG 199

Db 366 ACACGCAAGATGAGCGGGCGCTAGAGTGCCACCACTTTTGGCGCGTGGGAGATACAGT 425

Qy 200 ARACNATGAARGAGINTYNGARCNGCNGCNGTGGATNCCNYNGTNGATGAARCART 259

Db 426 GCTCCCGGACCTCAAGTTCTTCTGTGTAGATGATACAGCCCATCTGCTGGAGGACT 485

Qy 260 GYCAICNGAYACNAARAATYTYNTTYGWSNTYNTTYGNCNCNGTNGTYTNGAYGAY 319

Db 486 ACAAGAGCTCTGCGCCCTGTCGCTGTGTGTGAACGCGCAAGCGCGCTGCGCGC 545

Qy 320 TNGAYGARACNATNCAACCCNTGTCAYWSNTYNTGYTNCARGTNAARGAYMGTGCGNC 379

Db 546 CGCTCATGCGCCAGTAGCGCTTTCGTGTGCTGACCGCATGCGCTGCGTGTGCGCG 605

Qy 380 CNGTATGWSGNCNTYGGNTTYCCNTGCGCNGAYATGYTNGARTGYGAYMGTNTYCCNC 439

Db 606 AG 607

Qy 440 AR 441

RESULT 14

LOCUS HSU43318 2334 bp mRNA PRI 24-FEB-1996

DEFINITION Human putative transmembrane receptor (frizzled 5) mRNA, complete cds.

ACCESSION U43318

NID g1151251

KEYWORDS human.

SOURCE Homo sapiens

ORGANISM Homo sapiens

Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata; Vertebrata; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

1 (bases 1 to 2334)

Wang,Y., Macke,J.P., Abella,B.S., Andreasson,K., Worley,P., Gilbert,D.J., Copeland,N.G., Jenkins,N.A. and Nathans,J.

A large family of putative transmembrane receptors homologous to the product of the Drosophila tissue polarity gene frizzled

J. Biol. Chem. 271 (8), 4468-4476 (1996)

96224032

2 (bases 1 to 2334)

Abella,B., Wang,Y., Macke,J.P. and Nathans,J.

Direct Submission

Submitted (14-DEC-1995) Jeremy Nathans, Molecular Biology and Genetics, Johns Hopkins Medical School, 725 N. Wolfe Street, Baltimore, MD 21205

Location/Qualifiers

1. .2334

/organism="Homo sapiens"

/db_xref="taxon:9606"

/chromosome="2"

/map="2q33-34"

/tissue="retina"

321. .2078

/gene="frizzled 5"

321. .2078

/gene="frizzled 5"

/note="putative transmembrane receptor"

/codon_start=1

/product="transmembrane receptor"

/db_xref="PID:g1151252"

/translation="MARPDSPAPPSLLLLLQVLVGRRAAASKAPVQEIIVPMCRGI GYNLTHMPNQNHDTQDEAGLEVHGFVPLVEIQSPDLRFFLCMTYTCIDPHKPL PPCRVCERAKAGSPLMROYGFAPWDMFCDRLPVLRDAEVLCDYNRSEATAPP RPFAPKPTLPGPGAGGCGPCAGPCFVCRPFVILKESHPLYNKVRTGOVPC AVPCYOPSFADERTFATFGLKSLVLCFISTTVAFLDMDTFRYPERPIIFLSA CYLCSGLFVLYVGHASVACSEHNNHIVETGTPALCTVFLVLYFEFGNASSIWWV ILSLTFWLAAMKMGNEAIGYQGFHLAAGVLSVLRKSVIKQGGTKTDKLEKLMIRI GQNLNSLRFLVGLPLVLYLVLGTLFLAGVSLFRKSVIKQGGTKTDKLEKLMIRI GIFTLYTPASIVVACVLYEQHYRESWEAALTACPGHDTGQPRAPKEVWVLMKYF MCLVYGITSGVWWSGKTVESWRRTSRCCCRPRRGRHSGGMAAGDYPEASAALTGR TGPPGPAATYHKQVSLSHV"

BASE COUNT 356 a 803 c 736 g 439 t

ORIGIN

Query Match 10.2%; Score 75; DB 20; Length 2334;

Best Local Similarity 39.7%; Pred. No. 3.50e-14;

Matches 116; Conservative 44; Mismatches 132; Indels 0; Gaps 0;

Db 439 TGTGCGCGGCGATCGGCTACACCTGACGCACATGCCAACCAAGTTCAACACGACACC 498

Qy 146 TGTGCAYGAGNATGARTAYCARAAYATGMGNTNCCNAAYYNTYNGGNCAYGACNA 205

Db 499 AGCAGAGCGCGGCGCTGGAGGTGCACCACTTCTGCGCGCTGGAGATCCCAATGCTCC 558

Qy 206 TGAARGARTINTNGARCARGCNGGNGCTGGATNCCNTNGTNGATGAARCARTGYCATC 265

Db 559 CGGACATGCGCTTCTTCCTATGACTATGATACACGCCCATCTGTCTGCCCGACTACCA 618

Qy 266 CNGAYACNAARAATYTYNTTYGWSNTYNTTYGNCNCNTGTYTNGAYGAYTNGAYG 325

Db 619 AGCGGCTGCGCGCTGCGCTGCTGGAGCGGCGGCGGCGGCTGCTGCGCGCTGA 678

Search completed: Wed Aug 5 04:00:29 1998
Job time : 1099 secs.

THIS PAGE BLANK (USPTO)

Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

(TM)

MPSrch_nn n.a. - n.a. database search, using Smith-Waterman algorithm
Run on: Wed Aug 5 04:16:55 1998; MasPar time 140.38 Seconds
Tabular output not generated.
Title: >rev-trans-seq-2
Description: (1-885) from translate.seq
Perfect Score: 734
N.A. Sequence: 1 ATGTCNAGCGNCCGNCWS.....SNATMGNAAATTCARTGY 885
Comp: TACRANGTCCGNCNCNWS.....SNTAKNCNTTYRANGTYACR

Scoring table: TABLE default
Gap 6
Nmatch STD : Dbase 0; Query 0
Searched: 188442 seqs, 68026449 bases x 2
Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: n-geneseq32
1:part1 2:part2 3:part3 4:part4 5:part5 6:part6 7:part7
8:part8 9:part9 10:part10 11:part11 12:part12 13:part13
14:part14 15:part15 16:part16 17:part17 18:part18
19:part19 20:part20 21:part21 22:part22 23:part23
24:part24 25:part25 26:part26 27:part27 28:part28
29:part29 30:part30 31:part31 32:part32 33:part33
34:part34 35:part35 36:part36 37:part37 38:part38
39:part39 40:part40

Statistics: Mean 11.324; Variance 13.136; scale 0.862
Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES			
Result No.	Score	Query Match	Pred. No.
1	89	12.1	2421 37
2	75	10.2	2334 37
3	64	8.7	1770 37
4	63	8.6	2828 37
5	61	8.3	2344 37
6	59	8.0	501 3
7	59	8.0	2259 37
8	56	7.6	501 3
9	54	7.4	501 3
10	53	7.2	498 3
11	53	7.2	501 3
12	53	7.2	565 6
13	52	7.1	501 3

14	51	6.9	501 3	N50023	Sequence encoding new	4.41e-05
15	51	6.9	501 3	N50032	Sequence encoding new	4.41e-05
16	50	6.8	501 3	N50023	Sequence encoding new	8.82e-05
17	49	6.7	501 3	N50024	Sequence encoding new	1.76e-04
18	49	6.7	501 3	N50024	Sequence encoding new	1.76e-04
19	49	6.7	984 17	Q94336	Degenerate Alteromona	1.76e-04
20	48	6.5	501 3	N50026	Sequence encoding new	3.49e-04
21	48	6.5	501 3	N50027	Sequence encoding new	3.49e-04
22	47	6.4	498 3	N50034	Sequence encoding new	6.91e-04
23	47	6.4	501 3	N50029	Sequence encoding new	6.91e-04
24	47	6.4	501 3	N50031	Sequence encoding new	6.91e-04
25	47	6.4	501 3	N50032	Sequence encoding new	6.91e-04
26	47	6.4	501 3	N50028	Sequence encoding new	6.91e-04
27	46	6.3	1302 38	T90224	Magnetospirillum sp.	1.36e-03
28	46	6.3	2624 37	T89886	Mouse frizzled gene 3	1.36e-03
29	45	6.1	501 3	N50030	Sequence encoding new	2.88e-03
30	45	6.1	501 3	N50033	Sequence encoding new	2.88e-03
31	45	6.1	501 3	N50027	Sequence encoding new	2.88e-03
32	44	6.0	501 3	N50024	Sequence encoding new	5.24e-03
33	43	5.9	501 3	N50026	Sequence encoding new	1.02e-02
34	42	5.7	172 32	T76363	Human interleukin 8 a	1.98e-02
35	42	5.7	204 1	N81164	Base substituted E.co	1.98e-02
36	41	5.6	204 1	N81164	Base substituted E.co	3.83e-02
37	41	5.6	565 6	Q35072	HCV envelope region n	3.83e-02
38	41	5.6	657 7	Q43519	Degenerate FMN reduct	3.83e-02
39	41	5.6	1065 17	Q94335	Degenerate Alteromona	3.83e-02
40	40	5.4	91 9	O51746	Oligonucleotide probe	7.36e-02
41	40	5.4	2478 31	T67287	Soluble starch synthase	7.36e-02
42	40	5.4	3871 2	N71302	HSV-1 gB and surround	7.36e-02
43	39	5.3	91 9	O51746	Oligonucleotide probe	1.41e-01
44	39	5.3	501 3	N50025	Sequence encoding new	1.41e-01
45	39	5.3	984 17	Q94336	Degenerate Alteromona	1.41e-01

ALIGNMENTS

RESULT 1
ID T89892 standard; DNA; 2421 BP.
AC T89892:
DT 27-APR-1998 (first entry)
DE Mouse frizzled gene 8 (Mfz8) encoding a Wnt receptor.
KW Wnt receptor; mouse frizzled gene 8; Mfz8 gene;
KW signal transduction; cancer; cell growth; cell proliferation;
KW mammary tumour; oncogene; therapy; ds.
OS Mus musculus.
FH Key Location/Qualifiers
FT CDS 188..2245
FT /*tag= a
PN WO9739357-A1.
PD 23-OCT-1997.
PF 11-APR-1997; U06049.
PR 12-APR-1996; US-015307.
PA (UJO) UNIV JOHNS HOPKINS.
PA (STRD) UNIV LELAND STANFORD JUNIOR.
PI Andrew D, Bhanot P, Brink M, Hsieh J, Nathans J,
PI Nusse R, Samos CH, Wangy;
DR WPI: 97-526631/48.
DR P-PSDB: W31274.
PT Identification of Wnt receptor binding modulators - useful for
PT treatment of cancer and growth, development or proliferation related
PT disorders
PS Disclosure: Page 46-47; 61pp; English.
CC This sequence comprises novel mouse frizzled gene 8 (Mfz8)
CC that encodes a transmembrane receptor, frizzled-8 (see W31274),
CC a Wnt receptor (WntR). Novel frizzled family members have been
CC identified in Drosophila, mouse, human and Caenorhabditis (see
CC T89885-92) and are considered to be Wnt receptors. Wnt receptors
CC can be used in a novel, claimed method of screening for compounds
CC which modulate the binding of a Wnt polypeptide (secreted proteins
CC involved in cell-to-cell signalling) to a Wnt receptor. Wnt is
CC involved in (mammary) cancer and other processes involving growth,
CC development and proliferation (both normal and abnormal).
CC Modulators identified by the claimed method are useful for

Db 310 garaaytbataargargayttvacnmddggnnaarytbatgwnnnwnnytbcaaytbaar 369
 Cp 283 RRAAYTTTNGTTCNGGRTGRCAYTGYTTTCATNACNARNNGNATCCANGCNCNGCYT 224
 Db 370 mgdtayaygngmdathytbcaytaytbaargcnaargartaywnncaytgygcntgg 429
 Cp 223 GYTCNARNACYTCYTTCATNGTTCRTGNCNARNARRTTNGGNARCKCATRTTGTGT 164
 Db 430 acnathgmdgntngarathytbmgaaytatyttaytatythaaymgdytbaacngntay 489
 Cp 163 AYTCAATNCCRTGRCANARYTGNARRTTNGCNGNATNGGTYTTCARTTNSWCKYTTT 104
 Db 490 ytbmgdaay 498
 Cp 103 ANSWAART 95

RESULT 10
 ID N50034 standard; DNA; 498 BP.
 AC N50034;
 DT 04-SEP-1991 (first entry)
 DE Sequence encoding new modified human beta interferon polypeptides
 DE IFNX 485.
 KW Antiviral; cell growth regulator; immune system regulator;
 KW antiproliferative; ss.
 OS Homo sapiens.
 FH Key Location/Qualifiers
 FT cds 1..498
 FT /*tag= a
 PN EP-163993-A.
 PD 11-DEC-1985.
 PF 17-MAY-1985; 105750.
 PR 17-MAY-1984; GB-012564.
 PA (SEAR) SEARLE G D & CO.
 PI Bell LD, Boseley PG, Porter AG;
 DR WPI: 85-311944/50.
 DR P-PSDB: P50033.
 PT New modified human beta interferon polypeptide(s) - prepd. by
 PT plasmid transformed bacteria, with improved antiviral,
 PT anti-proliferative and immune regulating actions
 PS Claim 28; Chart 21, page 43; 71pp; English.
 CC Compared with interferon beta prep. by recombinant methods, the
 CC INFs of the invention are more active and have different affinities
 CC for cell surface receptors (allowing selective targeting); they
 CC have higher therapeutic index; improved stability against microbial
 CC breakdown during synthesis; and better in vivo solubility and
 CC stability. They are also easier to recover from incubation mixts.
 SQ Sequence 498 BP; 112 A; 30 C; 68 G; 77 T;

Query Match 7.2%; Score 53; DB 3; Length 498;
 Best Local Similarity 7.8%; Pred. No. 1.09e-05;
 Matches 38; Conservative 183; Mismatches 264; Indels 2; Gaps 2;
 Db 10 aaytytbgngnttytbcarmgdnwnnaayttycarwnncaraarytbytbgtgcar 69
 Cp 580 ARTCTTYTTCANARNGTYTCCATNATRTCTTTCRTCTCTCTTTTTRTCANG 521
 Db 70 ytbbaaygngmdybtbartaytytbaargaymgdgcngaytayaathccnatggar 129
 Cp 520 CYTCRCANACYTTTNGGNGCYTCYTCNCTNGCNGGNARNTGTCNSWSNGCNARNG 461
 Db 130 atg-acngaraargarttycaraargaygngcngnytbacnathaygaratgytba 188
 Cp 460 GNATRCANARTCTTTTCYTGNGGRANCKRTCCAYTCN-ARCATRTCTNGGCCANGR 402
 Db 189 raayathtygcnathtymgdcaayaywnnnwnnnacngntggaygaracnathgt 248
 Cp 401 AANCCRAANGCNSWCATNACNGGRCRCANCKRTCYTTNACYTGNACRCANARNSWRTGR 342
 Db 249 ngaraaytytbgcnaaygntntaycaycarathaaaycaytbaaracngntnytbarga 308
 Cp 341 CANGGYTGNTATNGTTCNARTCTCNARRCANACNACNGGRCRAANARNSWRCANARR 282

Db 309 raarytbgaraargayttvacnmddggnnaarytbatgwnnnwnnytbcaaytbaarmg 368
 Cp 281 AAYTTTNGTTCNGGRTGRCAYTGYTTTCATNACNARNNGNATCCANGCNCNGCYTGY 222
 Db 369 dtaaytaygngmdathytbcaytaytbaargcnaargartaywnncaytgygcntggac 428
 Cp 221 TCNARNACYTCYTTCATNGTTCRTGNCNARNARRTTNGGNARCKCATRTTGTGTAY 162
 Db 429 nathgtnmgdgtngarathytbmgaaytatyttaytatythaaymgdytbaacngntayt 488
 Cp 161 TCNATNCCRTGRCANARYTGNARRTTNGCNGNATNGGTYTTCARTTNSWCKYTTTAN 102
 Db 489 bmgdaay 495
 Cp 101 SWAART 95

RESULT 11
 ID N50028 standard; DNA; 501 BP.
 AC N50028;
 DT 04-SEP-1991 (first entry)
 DE Sequence encoding new modified human beta interferon polypeptides
 DE IFNX 445.
 KW Antiviral; cell growth regulator; immune system regulator;
 KW antiproliferative; ss.
 OS Homo sapiens.
 FH Key Location/Qualifiers
 FT cds 1..501
 FT /*tag= a
 PN EP-163993-A.
 PD 11-DEC-1985.
 PF 17-MAY-1985; 105750.
 PR 17-MAY-1984; GB-012564.
 PA (SEAR) SEARLE G D & CO.
 PI Bell LD, Boseley PG, Porter AG;
 DR WPI: 85-311944/50.
 DR P-PSDB: P50027.
 PT New modified human beta interferon polypeptide(s) - prepd. by
 PT plasmid transformed bacteria, with improved antiviral,
 PT anti-proliferative and immune regulating actions
 PS Claim 28; Chart 2f, page 37; 71pp; English.
 CC Compared with interferon beta prep. by recombinant methods, the
 CC INFs of the invention are more active and have different affinities
 CC for cell surface receptors (allowing selective targeting); they
 CC have higher therapeutic index; improved stability against microbial
 CC breakdown during synthesis; and better in vivo solubility and
 CC stability. They are also easier to recover from incubation mixts.
 SQ Sequence 501 BP; 114 A; 31 C; 68 G; 79 T;

Query Match 7.2%; Score 53; DB 3; Length 501;
 Best Local Similarity 7.7%; Pred. No. 1.09e-05;
 Matches 36; Conservative 180; Mismatches 247; Indels 3; Gaps 2;
 Db 33 dwnnnnaayttycartgycaraatybtbtbggcarytbaayggmgdytbgartayt 92
 Cp 557 ATNATCTCTTCRTCTCTCTTTTTCATNGCYTCRCANACYTTTNGGNGCYTCY 498
 Db 93 ytbbaargaymgdatgaaytytgayathcngcngargathaaarcarcngcarctyca 152
 Cp 497 TCNGTNGCNGGNARNARRTGTCNNSWSNGCNARNGNATRCANARRCTTTR--TCYT 440
 Db 153 raargaygngcngnytbacnathaygaratgytbcaraayathtygcnathtymg 212
 Cp 439 GNGGRAANCKRTCCAYTCN-ARCATRTCTNGGCCANGGRAANCCRAANGCNSWCATNACN 381
 Db 213 dcargaywnnnwnnnacngntggaygaracnathgtngaraaytytbytcnnaaygt 272
 Cp 380 GNGGRCANCKRTCYTTNACYTGNACRCANARNSWRTGRCANGTYTGNATNGTYTCTCN 321
 Db 273 ntaycaycarathaaaycaytbaaracngntnytbggargaraarytbgaraargaytt 332
 Cp 320 ARTCTCNARRCANACNACNGGRCRAANARNSWRCANARRAAYTYTITNGTTCNGGRTGR 261

RESULT 14

ID N50023 standard; DNA: 501 BP.
 AC N50032;
 DT 04-SEP-1991 (first entry)
 DE Sequence encoding new modified human beta interferon polypeptides
 DE IFNX 416.
 KW Antiviral; cell growth regulator; immune system regulator;
 KW antiproliferative; ss.
 OS Homo sapiens.
 FH Key Location/Qualifiers
 FT cds 1..501 /*tag= a
 FT EP-163993-A.
 PN 11-DEC-1985.
 PD 17-MAY-1985; 105750.
 PR 17-MAY-1984; GB-012564.
 PA (SEAR) SEARLE G D & CO.
 PI Bell LD, Boseley PG, Porter AG;
 DR WPI: 85-311944/50.
 DR P-PSDB: P50022.
 PT New modified human beta interferon polypeptide(s) - prepd. by
 PT plasmid transformed bacteria, with improved antiviral,
 PT anti-proliferative and immune regulating actions
 PS Claim 28; Chart 2a, page 32; 71pp; English.
 CC Compared with interferon beta prepd. by recombinant methods, the
 CC INFs of the invention are more active and have different affinities
 CC for cell surface receptors (allowing selective targeting); they
 CC have higher therapeutic index; improved stability against microbial
 CC breakdown during synthesis; and better in vivo solubility and
 CC stability. They are also easier to recover from incubation mixts.
 CC Sequence 501 BP; 107 A; 31 C; 69 G; 80 T;
 SQ
 Query Match 6.9%; Score 51; DB 3; Length 501;
 Best Local Similarity 16.0%; Pred. No. 4.41e-05;
 Matches 50; Conservative 95; Mismatches 167; Indels 1; Gaps 1;
 Db 166 gcnvthacnathaygaratgctbcaaaathttgcnathttmgdcargaywnnnwn 225
 QY 470 SNWSGAYCAIYNYINCCNGCWACNGARGCNCNCNARGNTGYGARGNTGYAARA 529
 Db 226 wnnacngntggaaygaracnathgtngaraaytbythbgcnaaygtntaycaycarat 284
 QY 530 AYAAARAAYGAYGAYAAAYGATATNATGARGACNTINTGYAARAAYGAYTYGCTNA 589
 Db 285 haaycsythaaracngntybgatgataaarytbgaraargaytthyacnmgdgnaa 344
 QY 590 ARATNAARGTNAARGARATNACNTAYATNAAYMGNGAYACNARATNATNTNGARACNA 649
 Db 345 rytbatgwnnnnythbcaytbaarmgdtaytggngmdathytbcaytavytbaargc 404
 QY 650 ARWSNARACNATNAYARIYNAAYTGGNGTWSNGARNGAYTNAARAARWSNGTNY 709
 Db 405 naargattaywnncaytgygcntggacnathgtngmdgtngarathytbmgdtaaytca 464
 QY 710 TNGGYTNAARGAYWSNYINCACTGYACNTGYGARGARATGAAYGAYATNAAYGCNCNT 769
 Db 465 yttaythaaymgd 477
 QY 770 AYTNGTNGATGG 782

RESULT 15

ID N50032 standard; DNA: 501 BP.
 AC N50032;
 DT 04-SEP-1991 (first entry)
 DE Sequence encoding new modified human beta interferon polypeptides
 DE IFNX 449.
 KW Antiviral; cell growth regulator; immune system regulator;
 KW antiproliferative; ss.
 OS Homo sapiens.
 FH Key Location/Qualifiers

FT cds 1..501 /*tag= a
 FT EP-163993-A.
 PN 11-DEC-1985.
 PD 17-MAY-1985; 105750.
 PR 17-MAY-1984; GB-012564.
 PA (SEAR) SEARLE G D & CO.
 PI Bell LD, Boseley PG, Porter AG;
 DR WPI: 85-311944/50.
 DR P-PSDB: P50031.
 PT New modified human beta interferon polypeptide(s) - prepd. by
 PT plasmid transformed bacteria, with improved antiviral,
 PT anti-proliferative and immune regulating actions
 PS Claim 28; Chart 2j, page 41; 71pp; English.
 CC Compared with interferon beta prepd. by recombinant methods, the
 CC INFs of the invention are more active and have different affinities
 CC for cell surface receptors (allowing selective targeting); they
 CC have higher therapeutic index; improved stability against microbial
 CC breakdown during synthesis; and better in vivo solubility and
 CC stability. They are also easier to recover from incubation mixts.
 CC Sequence 501 BP; 108 A; 30 C; 72 G; 79 T;
 SQ

Query Match 6.9%; Score 51; DB 3; Length 501;
 Best Local Similarity 7.6%; Pred. No. 4.41e-05;
 Matches 37; Conservative 187; Mismatches 262; Indels 3; Gaps 2;
 Db 10 aaytbytbgnnttytbcarmgdnwnnnaayttcartyccarayaarbytbgtgcar 69
 Cp 580 ARTCTTTTTCANARNNTYTCATNATRTCTTTCRTCTCTTCTTCTTCTTCTTCTTCT 521
 Db 70 ybaaygngmdybtgartytgytbaargaymgdcaytgytggnttyccngnggar 129
 Cp 520 CYTCRCANACYTTNGGNGCYTCYTCNGTNGCNGGNARNARNARTGRTCSNWSNNGCNA 461
 Db 130 garttygaygnaaycarttycaraargaygcngcnytbacnathaygaratgyt 189
 Cp 460 GNATRCANARTCTRTT--TCYTNGGGAACNCKRTCTTCAYTCN-ARCATTCNGGCCANG 404
 Db 130 caraayathtygcgnathtymgdcargaywnnnwnnnacngnttggaaygaracnath 249
 Cp 403 GRAANCCRAANGCNSWCATNACNGGNGCNCACNCKRTCTTTCAYTCNACRCANARNSWRT 344
 Db 250 gtngaraaytbythbgcnaaygtntaycaycarathaaycaytbaaracngntybgar 309
 Cp 343 GRCANGGYTGNATNGTYTCRCNARTCTCNARRCTCNARRCANACNGGNGCRAANARNSWRCANA 284
 Db 310 garaarytbgaraargaytthyacnmgdgnaarytbatgwnnnnythbcaytbaar 369
 Cp 283 RRAATTTTNGTTCNGGRTGRCAYTCYTTCATNACNARNGGNATCCANGCNCNGCYT 224
 Db 370 mgdtaytvgngmdathytbcaytavytbaargcnaaargartaywnncaytgygcntgg 429
 Cp 223 GYTCNARNACYTCYTTCATNGTYTCTCGCCNARNARNRTTNGGNARNCCKCATRTTGT 164
 Db 430 acnathgntmgdgtngaratybtbmgdtaaytthyathaaymgdytbacngntay 489
 Cp 163 AYTCNATCCCTGRCANARYTGNARFTTNGCNGGNATNGGTYTTCARTNSWNCKYTTT 104
 Db 490 ytbmgdaay 498
 Cp 103 ANSWRAART 95

Search completed: Wed Aug 5 04:19:20 1998
 Job time : 145 secs.

(TM)

Release 3.1A John F. Collins, Blocomputing Research Unit.

Copyright (C) 1993-1998 University of Edinburgh, U.K.

Distribution rights by Oxford Molecular Ltd

MPsrch_nn n.a. - n.a. database search, using Smith-Waterman algorithm

Run on: Wed Aug 5 04:00:48 1998; MasPar time 937.53 Seconds
 Tabular output not generated. 1322.069 Million cell updates/sec

Title: >rev-trans-seq-2

Description: (1-885) from translate.seq

Perfect Score: 734

N.A. Sequence: 1 ATGYNCARGGCCGCGGWS.....SNATWGNARNTNCARTGY 885

Comp: TACRANGTYCCGGGCCGWS.....SNTANKNTTYRANGTYACR

Scoring table: TABLE default

Gap 6

Nmatch STD : Dbase 0; Query 0

Searched: 1832099 seqs, 700269816 bases x 2

Post-processing: Minimum Match 0%
 Listing first 45 summaries

Database: emb1_est55

Database: 1:em_est10 2:em_est11

3:gb_est1 4:gb_est10 5:gb_est11 6:gb_est12 7:gb_est13
 8:gb_est14 9:gb_est15 10:gb_est16 11:gb_est17
 12:gb_est18 13:gb_est19 14:gb_est2 15:gb_est20
 16:gb_est21 17:gb_est22 18:gb_est23 19:gb_est3
 20:gb_est4 21:gb_est5 22:gb_est6 23:gb_est7 24:gb_est8
 25:gb_est9 26:gb_gss

Statistics: Mean 15.281; Variance 4.528; scale 3.375

Pred. No. is the number of results predicted by chance to have a
 score greater than or equal to the score of the result being printed,
 and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Length	ID	Description	Pred. No.
1	357	48.6	458 24	AA105749 ml84a03.r1 Stratagene	0.00e+00
2	345	47.0	482 23	W77968 zd70e11.r1 Soares feta	0.00e+00
3	319	43.5	396 25	AA073862 mm97f07.r1 Stratagene	0.00e+00
4	304	41.4	426 20	H87071 ys74d07.r1 Homo sapien	0.00e+00
5	303	41.3	378 10	AA429960 zw58g10.r1 Soares tota	0.00e+00
6	294	40.1	359 4	AA295018 EST100474 Pancreas tum	0.00e+00
7	279	38.0	406 19	H16121 y128c05.r1 Homo sapien	0.00e+00
8	277	37.7	414 22	AA024771 ze76g10.r1 Soares feta	0.00e+00
9	244	33.2	302 4	AA330758 EST34493 Embryo, 6 wee	8.91e-270
10	242	33.0	344 19	H14917 y126d03.r1 Homo sapien	5.02e-267
11	235	32.0	394 25	AA260087 va52g08.r1 Soares mous	2.08e-257
12	210	28.6	299 22	W08345 mb41f02.r1 Soares mous	3.11e-223
13	193	26.3	248 4	AA297695 EST113252 Fetal heart	3.62e-200

14	176	24.0	219	4	AA297550	EST113078 Fetal kidney	2.85e-177
15	165	22.5	562	7	AA449032	zx07e10.r1 Soares tota	1.45e-162
16	163	22.2	231	19	H29095	ym31e03.r1 Homo sapien	6.71e-160
17	151	20.6	250	23	AA061047	mj78a12.r1 Soares mous	5.48e-144
18	148	20.2	184	24	AA209833	mo90h01.r1 Beddington	4.97e-140
19	145	19.8	523	25	AA122822	mr03c11.r1 Soares mous	4.41e-136
20	144	19.6	188	7	AA457909	vf74b03.r1 Soares mous	9.09e-135
21	143	19.5	477	17	AA934293	ka07g11.s1 Onchocerca	1.87e-133
22	140	19.1	482	18	AA922948	oh10h11.s1 NCI CGAP Co	1.60e-129
23	136	18.5	287	24	AA082155	ze88f06.r1 Soares feta	2.70e-124
24	132	18.0	518	17	AA527991	on47d08.s1 NCI CGAP Co	4.37e-119
25	128	17.4	334	17	AA834426	of68d09.r1 NCI CGAP Co	6.75e-114
26	109	14.9	134	4	AA295743	EST101187 Thymus III H	1.44e-99
27	98	13.4	501	10	AA424593	zv91h11.s1 Soares NHM	8.98e-76
28	92	12.5	520	10	AA424647	zv91h11.r1 Soares NHM	2.32e-68
29	91	12.4	438	10	AA424594	zv91h12.s1 Soares NHM	3.90e-67
30	86	11.7	542	13	AA775798	ad14g08.s1 Soares NBHF	4.83e-61
31	81	11.0	252	13	AA754459	97SN1787 Rice Immature	5.04e-55
32	76	10.4	252	13	AA754459	97SN1787 Rice Immature	4.35e-49
33	72	9.8	337	16	AA907431	o102f01.s1 NCI CGAP Lu	2.09e-44
34	72	9.8	369	10	AA423653	ve80h07.r1 Soares mous	2.09e-44
35	66	9.0	424	8	AA481448	zv45b07.s1 Soares ovar	1.63e-37
36	65	8.9	276	5	AA371169	EST83193 Prostate gnan	2.21e-36
37	64	8.7	247	13	AA754458	97SN1784 Rice Immature	2.97e-35
38	63	8.6	531	3	M89402	CEL08B2 Caenorhabditis	3.93e-34
39	61	8.3	247	13	AA754458	97SN1784 Rice Immature	6.63e-32
40	61	8.3	248	5	AA375443	EST87993 HSC172 cells	6.63e-32
41	61	8.3	502	24	AA020088	mh49d08.r1 Soares mous	6.63e-32
42	60	8.2	376	9	C39238	C.elegans cdna clone y	8.44e-31
43	58	7.9	401	13	AA724275	ah91c05.s1 Soares NFL	1.31e-28
44	57	7.8	360	6	C09584	C.elegans cdna clone y	1.60e-27
45	56	7.6	360	9	C46856	C.elegans cdna clone y	1.92e-26

ALIGNMENTS

RESULT 1
 LOCUS AA105749 458 bp mRNA EST 04-FEB-1997
 DEFINITION ml84a03.r1 Stratagene mouse kidney (#937315) Mus musculus CDNA
 Clone 518668 5' similar to TR.G1151260 G1151260 TRANSMEMBRANE
 RECEPTOR. ; mRNA sequence.

ACCESSION AA105749
 NID G1654838
 KEYWORDS EST.
 SOURCE house mouse.
 ORGANISM Mus musculus
 Eukaryote; mitochondrial eukaryotes; Metazoa; Chordata;
 Vertebrata; Euthera; Rodentia; Sciurognathi; Muridae; Murinae;
 Mus.

REFERENCE 1 (bases 1 to 458)
 AUTHORS Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
 Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
 Scheinberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
 Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
 Waterston,R.
 TITLE The WashU-HMI Mouse EST Project
 JOURNAL Unpublished (1996)
 COMMENT Contact: Marra M/Mouse EST Project
 WashU-HMI Mouse EST Project
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: mouseest@watson.wustl.edu
 This clone is available royalty-free through LILNL; contact the
 IMAGE Consortium (info@image.lilnl.gov) for further information.
 MGI:312516
 Seq primer: -28ml3 rev1 ET from Amersham
 High quality sequence stop: 457.
 Location/Qualifiers
 1..458
 /organism="Mus musculus"

FEATURES
 source

```

/strain="C57/Bl6"
/note="vector: plasmid SK-; Site_1: EcoRI; Site_2:
XhoI; Cloned unidirectionally. Primer: Oligo dT. Average
insert size: 1.0 kb; Uni-ZAP XR vector; -5' adaptor
sequence: 5' GAATTCGCGACGAG 3' -3' adaptor sequence: 5'
CTGAGTGTGTTTTTTTTTTT 3'
/db_xref="taxon:10090"
/clone="518668"
/clone_lib="Stratagene mouse kidney (#937315)"
/sex="females"
/dev_stage="4 weeks"
/lab_host="SOLR (kanamycin resistant)"
<1. >458
BASE COUNT      119 a  140 c  116 g   83 t
ORIGIN

Query Match      48.6%; Score 357; DB 24; Length 458;
Best Local Similarity 60.6%; Pred. No. 0.00e+00;
Matches 280; Conservative 103; Mismatches 75; Indels 4; Gaps 4;

Db 1 AGCAGCGCGCATGGATTCCGCTGGTCATGACCACTGCCACCGCACACCAAGAGT 60
Qy 221 ARCARGCGGNCNTGGATNCNNTYNGTNGTGAARCACTGTCATCCNGAYACNARAART 280
Db 61 TCCTGTGCTGCTTCCGCCCTGTCTCTCGAGCACTAGATGAGACCATCCAGCCGT 120
Qy 281 TTYTNGTWSNYTNTTYGCNCCNGTGTGTYNGAYGAYTNGAYGARACNATNCARCNT 340
Db 121 GTCACCTGCTCTG-GTGAGGTGAAGACCGCTGCGCCCGCTCATGTCGCCCTTCG-CT 178
Qy 341 GYCAWNSYNTGTGTCNARGTNAARGAYMGNTGYCCNCCNGTNGTWSNCCNTTYGNT 400
Db 179 TCCCTGCGCCACATGCTGGAGTGGACCGCTTCCCGCAG-ACAACGACCTCTGCATCC 237
Qy 401 TYCCNTGGCCGAYATGTYNGARTGYGAYMGNTTYCCNARGAYAYGYTNGTYATNC 460
Db 238 CCCTCGCTAGTAG-GACCACCTCTCCCGGCGCACAGAGCAAGCTCCCAAGGTGTGTAAG 296
Qy 461 CNYTNGCWNWSNGAYCAYTYNTNCCNGCNACNGARGCNCNNAARCTNGYGARG 520
Db 297 CCTGCAAAACCAAGATGAGGACGACACGACATCATCGGAACCCCTTGTAAAATGACT 356
Qy 521 CNTGYAARAAYAAAYGAYGAYAYGATYATNATGABACNATYNTGYAARAAYGAYT 580
Db 357 TCGCACTGAAATCAAGTGAAGGATACGTCATCAACAGACAGACACCAAGATCATCC 416
Qy 581 TYGCNTNAARATNARGTNAARGARATNACNTAYATNAAYNGAYACNARATNATNY 640
Db 417 TGGAGACAAAGCAGACCACTTTACAAGCTGAACGGCGTGT 458
Qy 641 TNGARACNAARWSNAARACNATNTAYARYTNAAYGGNGTNG 682

RESULT 2
LOCUS      W77968      462 bp      mRNA      EST      17-OCT-1996
DEFINITION zdf0ell.r1 Soares fetal heart NBH19W Homo sapiens cDNA clone
sequence: 5', mRNA sequence.
ACCESSION  W77968
NID         61388502
KEYWORDS    EST.
SOURCE      human.
ORGANISM    Homo sapiens
Eukaryotes; mitochondrial eukaryotes; Metazoa; Chordata;
Vertebrata; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 462)
AUTHORS    Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,
Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,
Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F.,
Trevisakis,E., Waterston,R., Williamson,A., Wohlmann,P. and
Wilson,R.
TITLE       The WashU-Merck EST Project
JOURNAL     Unpublished (1995)
COMMENT

```

```

Contact: Wilson RK
WashU-Merck EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 1550 Std Error: 0.00
Seq primer: mob.REGA+ET
High quality sequence stop: 424.
Location/Qualifiers
1. 462
/organism="Homo sapiens"
/note="Organ: heart; Vector: pT7T3D (Pharmacia) with a
modified polylinker; Site_1: Not I; Site_2: Eco RI; lstr
strand cDNA was primed with a Not I - oligo(dT) primer [5',
TGTTCACATCTGAAGTGGAGCGCCGATCTTTTTTTTTTTT 3'],
double-stranded cDNA was size selected, ligated to Eco RI
adapters (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of a modified pT7T3 vector
(Pharmacia). Library went through one round of
normalization to a Cot = 5. Library constructed by
M.Fatima Bonaldo. This library was constructed from the
same fetus as the fetal lung library, Soares fetal lung
NBH19W."
/db_xref="taxon:9606"
/clone="346028"
/clone_lib="Soares fetal heart NBH19W"
/sex="unknown"
/dev_stage="19 weeks"
/lab_host="DH10B (ampicillin resistant)"
<1. >462
BASE COUNT      128 a  133 c  106 g   91 t   4 others
ORIGIN

Query Match      47.0%; Score 345; DB 23; Length 462;
Best Local Similarity 58.0%; Pred. No. 0.00e+00;
Matches 266; Conservative 108; Mismatches 81; Indels 4; Gaps 4;

Db 6 TCGCTCTTCGCCCGCTCTCGCTGATGATGACCTAGACGAGACCATCCAGCCACTCG 65
Qy 289 WSNYNTTYTTCNCCNGTNGTYTNGAYGAYTNGAYGARACNATNCARCCNTGYCAYWSN 348
Db 66 CTCTGN-TGCAGGTGAAGACCGCTGCGCCCGCTCATGTCGGCTTCGN-TTCCCTCGG 123
Qy 349 YTTGCTNCARGTNAARGAYMGNTGYCCNCCNGTNGTWSNCCNTTYGNTTYCCNTGG 408
Db 124 CCGACATGCTTGAGTCGACCGCTTCCCGCAGACACGACCTTTGCATCCCGCTCCGC 183
Qy 409 CCNGAYATGTYNGARTGYGAYMGNTTYCCNARGAYAYGYTNTGYATNCCNTNGCN 468
Db 184 TAGCAGNACCACCTCTCGCCAGCCGACGAGAGCTCCAAAGGTATGTGAAGCTGCAAA 243
Qy 469 WSNWSNGAYCAYTYNTNCCNGCNACNGARGCNCNNAARCTNGYGARGCNTGYAAR 528
Db 244 AATAAATGATGATGACACGACATATGAACGCTTTGTAAATGATTTTGCACCTG 303
Qy 529 AYAARAAYGAYGAYAYGATYATNATGARGCNTNTGYARAAYGAYTYGNTYN 588
Db 304 AAAATAAAGTGAAGAGATAACCTACATCAACCGAGATACCAAAATCATCTCGAAGAC 363
Qy 589 AARATNARGTNAARGARATNACNTAYATNAAYMGNGAYACNAAARATNATNGAR-AC 647
Db 364 CAAGAGCAAGACCATTTACAGCTGAACGGTGTGTCGCAAGGAGGACCTTGAAGAAATCGG 423
Qy 648 NAARWSNAARACNATNTAYARYTNAAYGGNGTNGWSNARGMNGAYTYN-AARAARWSNG 706
Db 424 TGCTGTGCTCAAGACAGCTTGCGAGTGCACCTGTGAAG 462
Qy 707 TNYTNGYTNAARGAYWSNTNCAATGYACNTGYGARG 745

```

```

3
RESULT 3
LOCUS AA073862 396 bp mRNA EST 15-FEB-1997
DEFINITION mm97f07.r1 Stratagene mouse heart (#937316) Mus musculus cDNA clone
536389 5' similar to FR:G151260 G151260 TRANSMEMBRANE RECEPTOR.
; mRNA sequence.
ACCESSION AA073862
NID g1595602
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryotes; mitochondrial eukaryotes; Metazoa; Chordata;
Vertebrata; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Mus.
REFERENCE 1 (bases 1 to 396)
AUTHORS Marra M., Hillier L., Allen M., Bowles M., Dietrich M., Dubuque T.,
Geisel S., Kucaba T., Lacy M., Le M., Martin J., Morris M.,
Schellenberg K., Streptoe M., Tan F., Underwood K., Moore B.,
Theising B., Wylie T., Lennon G., Soares B., Wilson R. and
Waterston R.
TITLE The WashU-HMI Mouse EST Project
JOURNAL Unpublished (1996)
COMMENT Contact: Marra M/Mouse EST Project
WashU-HMI Mouse EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@wustl.edu
This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:323325
Seq primer: -28ml3 rev1 ET from Amersham
High quality sequence stop: 390.
FEATURES
source
1..396
/organism="Mus musculus"
/strain="NIH/Swiss"
/Note="Vector: pluescript SK; Site_1: EcoRI; Site_2:
XhoI; Cloned unidirectionally. Primer: Oligo dr. 93
pooled NIH/Swiss 13 day embryo hearts. Average insert
size: 1.0 kb; Uni-ZAP XR vector; -5' adaptor sequence: 5'
GAATTCGCGCAGAG 3' -3' adaptor sequence: 5'
CTCGAGTCTTTTCTTTTCTTTT 3'"
/db_xref="taxon:10090"
/clone="536389"
/clone_lib="Stratagene mouse heart (#937316)"
/sex="pooled"
/dev_stage="13 day embryos"
/lab_host="SOLR (kanamycin resistant)"
<1..>396
93 a 126 g 103 g 74 t
mRNA
BASE COUNT 93 a 126 g 103 g 74 t
ORIGIN
Query Match 43.5%; Score 319; DB 25; Length 396;
Best Local Similarity 61.7%; Pred. No. 0.00e+00;
Matches 245; Conservative 88; Mismatches 62; Indels 2; Gaps 2;

401 TYCNTGCCNGAYATGYTNGARTGYGAYMGNTTYCCNCARGAAYAAAYGYYTNTGYATNC 460
239 CCCTCGCTAGTAGGACACCTCTCGCCGACAGGAGCTCCCAAGGTGTGTGAAG 298
461 CNYTNGCNSWNSGACAYATYNTNCCNGCAGARGAGCNCNARGTNTGYGARG 520
299 CCTGCAAAACCAAGATGAGGACGACACGACATCATCTGGAACCCCTTTGTAATAATGACT 358
521 CNTGYAARAAAYAAAYGAYGAYGAYATNATGGARACNTYNTGYAARAAYGAYT 580
359 TCGCAGTGAATCAAGTGAAGGAGATACGTACAT 395
581 TYGNTYNAARATNAAGTNAARGATNACNTATAT 617

4
RESULT 4
LOCUS H87071 426 bp mRNA EST 21-NOV-1995
DEFINITION Y874007.r1 Homo sapiens cDNA clone 220525 5' similar to SP:A45054
A45054 F2-1-PUTATIVE INTERCELLULAR SIGNAL TRANSDUCER OR TRANSMITTER
;
ACCESSION H87071
NID g1068650
KEYWORDS EST.
SOURCE human clone-220525 primer-M13RP1 library-Soares retina N2B4HR
vector-pT7T3D (Pharmacia) with a modified polylinker host-DH10B
(ampicillin resistant) Rsite1-Not I Rsite2-Eco RI 1st strand cDNA
was primed with a Not I - oligo(dT) primer
[5'-TGTTACCAATCTGAAGTGGAGCGCGCGCTTTTTTTTTTTT-3'],
double-stranded cDNA was size selected, ligated to Eco RI adapters
(Pharmacia), digested with Not I and cloned into the Not I and Eco
RI sites of a modified pT7T3 vector (Pharmacia). The retinas were
obtained from a 55 year old Caucasian male and total cellular
poly(A)+ RNA was extracted 6 hrs after their removal. The retina
RNA was kindly provided by Roderick R. McInnes M.D. Ph.D. from the
University of Toronto. Library constructed by Bento Soares and
M.Fatima Bonaldo.
ORGANISM Homo sapiens
Eukaryotes; Metazoa; Eumetazoa; Bilateria; Coelomata;
Deuterostomia; Chordata; Vertebrata; Gnathostomata; Osteichthyes;
Sarcopterygii; Choanata; Tetrapoda; Amniota; Mammalia; Theria;
Eutheria; Archonta; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 426)
AUTHORS Hillier L., Clark N., Dubuque T., Elliston K., Hawkins M.,
Holman M., Hultman M., Kucaba T., Le M., Lennon G., Marra M.,
Parsons J., Rifkin L., Rohlfing T., Soares M., Tan F.,
Trevaskis E., Waterston R., Williamson A., Wohlmann P. and
Wilson R.
TITLE The WashU-Merck EST Project
JOURNAL Unpublished (1995)
COMMENT Contact: Wilson RK
WashU-Merck EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@wustl.edu
High quality sequence stops: 355
Source: IMAGE Consortium, LLNL
This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
FEATURES
source
1..426
/organism="Homo sapiens"
/clone="220525"
<1..>426
96 a 128 c 110 g 83 t 9 others
mRNA
BASE COUNT 96 a 128 c 110 g 83 t 9 others
ORIGIN
Query Match 41.4%; Score 304; DB 20; Length 426;
Best Local Similarity 59.4%; Pred. No. 0.00e+00;
Matches 246; Conservative 91; Mismatches 73; Indels 4; Gaps 4;

```

```

Db 2  ACGCATNGAATACCAAGATCGGCTGCCACCTGCTGGCCACGAGACCATGAAGG 61
    ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 152 AYGCNATNGARTAYCARAAATGNGTNCNNAAYNYINGNCAYGARACNATGAAG 211
    ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 62  AGGTGCTGGAGAGCGCGCGCTGGATCCGCTGGTGCATGAAGCAGTGCACCCGGACA 121
    ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 212 ARGNTYNGARCARGCNGGNCNTGGATNCNNTNGTNGATGAARCARTGYCAYCCNGAYA 271
    ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 122 CCAAGAAGTTCCTGCTGCTCTCCGCCGCCGCTGCTGCTGATGACCTAGACGAGACCA 181
    ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 272 CNAARAARTYTYTNGWSNTNTTYGCNCCNGTNTGYTNGAYGAYTNGATGARACNA 331
    ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 182 TCCAGCCATGCACCTCGCTG-CTGCAGGTGAAAGACGCGCTGCGCCCGGTCATGCC 240
    ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 332 TNCARCCNTGYCAWSNTYGTGTCNARG-AYMGNITGYGCNCCNGATGWSN 390
    ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 241 GCCTTCGGTTCCTCCCTGCGCGACATGCTTGAGTGCAGCAGCTTCCCGCCAGCAAGCAG 300
    ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 391 GCNTTYGNTTYCCNTGCCNGAYATGTYTNGARTGYGAYMGNITTYCCNCARGAYAAYGAY 450
    ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 301 CTTTGCATCCCTCTGATGACGAGCAGCTTCTTCTGACGCCAGCAGCAAGTTNCAA 360
    ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 451 YNTGYATNCNNTYGCNWSNWSNGAYCAYTNTTNC-CNGCNACNARGARGCNCNAA 509
    ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 361 GGTATGTAAGCTGNAATAATTAATAATGATGATGAACAACGACATATGGGAA 414
    ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 510 RGTNTGYGRCNTGYAARAAYARAAYGAYGAYGA-YAAYGAYATNATGARA 562
    ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 5
LOCUS AA429960 378 bp mRNA EST 16-OCT-1997
DEFINITION zw58g10.r1 Soares total fetus Nb2HF8 9w Homo sapiens cDNA clone
W74306 5', mRNA sequence.
ACCESSION AA429960
NID G2113134
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryotae: Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 378)
AUTHORS Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisel,G., Jost,S.,
Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M., Martin,J.,
Moore,B., Schellenberg,K., Steptoe,M., Tan,F., Theising,B.,
White,Y., Wylie,T., Waterston,R. and Wilson,R.
WashU-Merck EST Project 1997
Unpublished (1997)
COMMENT Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Seq primer: -28m13 rev2 ET from Amersham
High quality sequence stop: 352.
FEATURES
source Location/Qualifiers
1..378
/organism="Homo sapiens"
/notes="Vector: pT73D-Pac (Pharmacia) with a modified
polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand
cDNA was prepared from mRNA obtained from pooled 8-9 week
(total) fetus material with a Not I - oligo(dT) primer [5'
TGTTCACTCAATGTAATGGAGCGCGCTTAAATTTTATTTTATTTT 3'].
Double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pT73 vector. Library
went through one round of normalization, and was
constructed by Bento Soares and M. Fatima Bonaldo."
/db_xref="taxon:9606"
/clone="774306"

```

```

/clone_lib="Soares total fetus Nb2HF8 9w"
/dev_stage="8-9 weeks"
/lab_host="DH10B"
BASE COUNT 115 a 101 c 88 g 74 t
ORIGIN
Query Match 41.3%; Score 303; DB 10; Length 378;
Best Local Similarity 59.7%; Pred. No. 0.00e+00;
Matches 222; Conservative 89; Mismatches 60; Indels 1; Gaps 1;
Db 8  CCGGTATGTCGCTTCG-CTTCCCTGCGCGACATGCTTGAGTGCAGCCTTCCCC 66
    ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 379 CCNNTNATGWSNNTTYGNTTYCCNTGCCNGAYATGTYTNGARTGYGAYMGNITTCN 438
    ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 67  CAGGACAAGACCTTTGCATCCCTGCTGACGAGCAGCAGCTCTCTGCGACCCAG 126
    ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 439 CARGAAYAGYATNTGYATNCCNTYGCNWSNWSNGAYCAYTNTTNCNCGNACNGAR 498
    ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 127 GAAGCTCCCAAGGTATGTGAAGCTGCAAAAATGATGATGACACGACATAATG 186
    ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 499 GARGCNCNARGTNTGYGARGCNTGYAARAAYARAAYGAYGAYGAYATNATG 558
    ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 187 GAACCTTTGTAATAATGATTTGCACGTGAATAAAGTGAAGGAGATACCTACATC 246
    ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 559 GARACNTYNTGYAARAAYGAYTTCNTYTAARATNAARGTNAARGARATNATAYATN 618
    ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 247 AACCGAGATACCAAAATCATCTGGAGACCAAGACGACCATTTTACAGCTGAACGCT 306
    ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 619 AAYMNGATACNARATNATNTNGARACNARWSNARACNATNTYARTTNAAYGN 678
    ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 307 GTGTCCGATAGGACCTGAAGAAATCGTGTGCTGCTCAAAAGACAGCTTGCACTCAC 366
    ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 679 GTNWSNGMNGAYTNAARAARWSNGTNTGTYTNAARGAYWSNTNCARTGYACN 738
    ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 367 TGTGAGGAGATG 378
    ||| ||||| |||||
Qy 739 TGYGARGARATG 750

RESULT 6
LOCUS AA295018 359 bp mRNA EST 18-APR-1997
DEFINITION EST100474 Pancreas tumor I Homo sapiens cDNA 5' end, mRNA sequence.
ACCESSION AA295018
NID g1947649
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryotae: mitochondrial eukaryotes; Metazoa; Chordata;
Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominidae;
Homo.
REFERENCE 1 (bases 1 to 359)
AUTHORS Adams,M.D., Kerlavage,A.R., Fleischmann,R.D., Fuldner,R.A.,
Bult,C.J., Lee,N.H., Kirkness,E.F., Weinstock,K.G., Gocayne,J.D.,
White,O., Sutton,G., Blake,J.A., Brandon,R.C., Man-Wai,C.,
Clayton,R.A., Cline,T.R., Cotton,M.D., Earle-Hughes,J., Fine,L.D.,
Fitzgerald,L.M., Fitchugh,W.M., Fritchman,J.L., Geoghagen,N.S.,
Glodek,A., Gnehm,C.L., Hanna,M.C., Hedblom,E., Hinkle,P.S., Jr.,
Kelley,J.M., Kelley,J.C., Liu,L.-I., Marmaros,S.M., Merrick,J.M.,
Moreno-Palauques,R.F., McDonald,L.A., Nguyen,D.T., Pelligrino,S.M.,
Phillips,C.A., Ryder,E., Scott,J.L., Saudek,D.M., Shirley,R.,
Small,K.V., Spriggs,T.A., Utterback,T.R., Weidman,J.F., Li,X.,
Bednarek,D.P., Cao,L., Cepeda,M.A., Coleman,T.A., Collins,E.J.,
Dimke,D., Feng,D.-F., Ferrie,A., Fischer,C., Hastings,G.A.,
He,W.W., Hu,J.S., Greene,J.M., Gruber,J., Hudson,P., Kim,A.K.,
Kozak,D.L., Kunsch,C., Hungjun,J., Li,H., Meissner,P.S., Olsen,H.,
Raymond,L., Wei,X.F., Wing,J., Xu,C., Yu,G.L., Ruben,S.M.,
Dillon,P.J., Fannon,M.R., Rosen,C.A., Haseltine,W.A., Fields,C.,
Fraser,C.M. and Venter,J.C.
Initial assessment of human gene diversity and expression patterns
based upon 83 million nucleotides of cDNA sequence
Nature 377 (6547 suppl), 3-174 (1995)
JOURNAL 96026280
MEDLINE
COMMENT Other_ESTs: THC180057

```

Contact: Kerlavage, AR
Bioinformatics
The Institute for Genomic Research
9712 Medical Center Drive, Rockville
Tel: 3018699056
Fax: 3018699423

Email: arkerlav@tiqr.org

EMBL: U01646v1;CC-0
For clone availability, additional sequence and expression
information related to this EST, please check the TIGR Human Gene
Index (<http://www.tigr.org/tdb/hgi.html>)
Seq primer: M13 Reverse.

Location/Qualifiers

1. 359

```

"organism":"Homo sapiens"
"note":"organ: pancreas; Vector: pBluescript SK-; Site:1:
ECOR1; Site:2: XhoI"
"db_xref":"ATCC (iphost):190769"
"db_xref":"taxon:9606"
"clone_lib":"pancreas tumor I"
"dev_stage":"adult"
"1_356"
"1_356"

```

BASE COUNT	106 a	96 c	82 g	68 t	7 others
ORIGIN					

Query Match 40.1%; Score 294; DB 4; Length 359;
Best Local Similarity 59.3%; Pred. No. 0.00e+00;
Matches 213: Conservative 84; Mismatches 62; Indels

Db	1	CGTGCAGGTGAAGACCGCTGCGCCCGCGGTCA	TGTCGGNNTCCCGCTGCGCCGA	60
QY	354	YGTNCARGTNAARGAYMGNTGYCNCNCGTNA	TGWSNGCNTYYGGNTYYCCNTGGCCGA	413
Db	61	CATGCTTGANTGCGACGCTTTCCCGCAGGACA	CAGACCTTTGCATCCCGCTCGCTAGCAG	120
QY	414	YATGYNGARTGYGAYMGNTTYCCNCARGAYA	YGAYYNTGYATNCCNTYNGCNWSNWS	473
Db	121	CGACCACCTCTCGCAGCCACCGAGGAGCTCA	AAAGGTATGTTNAGCGCTGCAGAAATAA	180
QY	474	NGAYCATYTTNCCNCGNACGARGCNCNARG	TNTGYGRCNGTGYAARAAYAA	533
Db	181	AANTGATGATGACAAACACATAATGGAACGC	TTGTAAATAATNTTTGCACTGAAAT	240
QY	534	RAAYGAYGAYAAAYGAYATNATGARACNY	TNTGYAARAAYGAYTYGTCNTNAARAT	593
Db	241	AAAAGTGAAGCAGATACCTACATCAACCGAG	ATACCAAATCATCTCGGAGACCAAGAG	300
QY	594	NAARGTNAARGARTNACNTAYATNAAYMG	NGAYACNAARATNTYTGARACNAARWS	653
Db	301	CAAGACCATTTACAAGCTGAACGGTGTGTCG	GAAGGACCGCTGAAGGAATCGGTGCTGT	359
QY	654	NAPACNATNTAAAYRTNAYGSGNTNWSGAR	BMGNVYTTNAAARAARNSGNTYNT	712

RESULT	7			
LOCUS	H16121	406 bp	mRNA	EST
DEFINITION	y128c05.r1 Homo sapiens cDNA clone 193560 5'.			
ACCESSION	H16121			
NID	9880941			
KEYWORDS				
EST.				

ORGANISM
Homo sapiens
and *M. ratina bonai*gou.
Eukaryotae; Metazoa; Eumetazoa; Bilateria; Coelomata;
Deuterostomia; Chordata; Vertebrata; Gnathostomata; Osteichthyes;
Sarcopterygii; Choanata; Tetrapoda; Amniota; Mammalia; Theria;
and *M. ratina bonai*gou.

REFERENCE
AUTHORS
 Eutheria: Archonta: Primates; Catarrhini: Hominoidea: Homo.
 1 (bases 1 to 406)
 Hillier, L., Clark, N., Dubuque, T., Elliston, K., Hawkins, M.,
 Holman, M., Hultman, M., Kucaba, T., Le, M., Lennon, G., Marra, M.,
 Parsons, J., Rifkin, R., Rohlfing, T., Soares, M., Tan, F.,
 Trevasaki, E., Waterston, R., Williamson, A., Wohlmann, P. and
 Wilson, R.
TITLE
 The WashU-Merck EST Project
JOURNAL
 Unpublished (1995)

Contact: Wilson RK
WashU-Merck EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810

FEATURES
source 1..406
IMAGE Consortium (info@image.lln
Location/Qualifiers
Source: IMAGE Consortium, LLNL
This clone is available royalty-
High quality sequence stops: 321
Small: 321

Query Match	38.08;	Score 279;	DB 19;	Length 406;
Best Local Similarity	58.94;	Pred. No. 0.00e+00;		
Matches	221;	Conservative	89;	Mismatches 60;
			Indels	5;
			Gaps	5;
Db	1	CCAGGACACGACCTTTGCATCCCTCGCTAGCAGCANCCACCCTCTCCAGCCACCG	60	
Qy	437	CNCARGAAYGAYYNTGYATNCCNTNGCNWSNNGAYCAYTNTNCCNGCNACG	496	
Db	61	AGGAAGCTCCAAAGGTATGTGAAGCCTGC AAAATATAAATATGATGACACAGACATAA	120	
Qy	497	ARGARGCNCNAAAGTNTGARGCNTGYAARAAYARAAYGAYGAYGAYAAYGAYATNA	556	
Db	121	TGGAACGCCTTCATAAAATGATTTCCTACTGAAATATAAAGTGAAGGAGATAACCTACA	180	
Qy	557	TGGARACNTNTGYAARAAYGAYTTTTCNTTNAARTNAAAGTNAARGATNACNTAYTA	616	
Db	181	TCAACCGAGATACCAAATCATCTCGAGACCAAGACAGACCATTTACAAGCTGAACG	240	
Qy	617	TNAAYMGAYACNAAATNATNTNGARACNAAARWSNAAACNATNTAYAARYTNAAYG	676	
Db	241	GTGTGTCGGAAGGACCTGAAGAAATCGGTGCTGTGGCTCAAGACAGCTTCGAGTGCA	300	
Qy	677	NGGTNWSNGARMNGAYTTNAARAARWSNGTNTNTGGYTNAARGAYWSNTNCACTGYA	736	
Db	301	CCTGTGAGGAGATGAACGACCATTCACGCGCCCTATCTGCTTCATGGGGACAGGAACAG	360	
Qy	737	CNTGARGARATGAAYGAYATN-AYGCMCCNTAYTNTGNTN-ATGGGN-CAR-AARCAR	792	
Db	361	GGTTGGGGAGCTGGT	375	
Qy	793	GGN-GGNGARYTNGT	806	

RESULT	8	AA024771	414 bp	mrna	EST	13-AUG-1996
LOCUS		ze76g10.r1	Soares fetal heart	NH0119W	Homo sapiens	CDNA clone
DEFINITION		364962 5',	mrna sequence.			
ACCESSION		AA024771				
NID		g1489677				
KEYWORDS		EST.				
SOURCE		human.				
ORGANISM		Homo sapiens				
		Eukaryotes; mitochondrial eukaryotes; Metazoa; Chordata;				
		Vertebrata; Eutheria; Primates; Catarrhini; Hominiidae; Homo.				
REFERENCE		1 (bases 1 to 414)				

AUTHORS

Hillier, L., Clark, N., Dubuque, T., Elliston, K., Hawkins, M., Holman, M., Hultman, M., Kucaba, T., Le, M., Lennon, G., Marra, M., Parsons, J., Rifkin, L., Rohlfing, T., Soares, M., Tan, F., Trevasakis, E., Waterston, R., Williamson, A., Wohlmann, P. and Wilson, R.

TITLE
JOURNAL
COMMENT

The WashU-Merck EST Project
Unpublished (1995)

Contact: Wilson RK

WashU-Merck EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108

Tel: 314 286 1800
Fax: 314 286 1810

Email: est@watson.wustl.edu

This clone is available royalty-free through LNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.

Seq primer: -28M13 rev2 from Amersham

High quality sequence stop: 288.

FEATURES

source

1. .414
/organism="Homo sapiens"
/note="Organ: heart; Vector: pT73D (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5', TGTTACATCTGAAGTGGAGCGGCATCTTTTTTTTTTTT 3'], double-stranded cDNA was size selected, ligated to Eco RI adapters (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of a modified pT73 vector (Pharmacia). Library went through one round of normalization to a Cot = 5. Library constructed by M.Fatima Bonaldo. This library was constructed from the same fetus as the fetal lung library, Soares fetal lung NHL19W.

/db_xref="taxon:9606"

/clone="364962"

/clone.lib="Soares fetal heart NbH19W"

/sex="unknown"

/dev_stage="19 weeks"

/lab_host="DHI0B (ampicillin resistant)"

<1. .>414

BASE COUNT 107 a 123 c 95 g 85 t 4 others

ORIGIN

Query Match 37.7%; Score 277; DB 22; Length 414;
Best Local Similarity 60.9%; Pred. No. 0.00e+00;
Matches 224; Conservative 85; Mismatches 54; Indels 5; Gaps 5;
Db 1 TGAAGCAGTCCACCGGACACCAAGTCTCTGCTCGCTCTTGGCCCGCTGTC 60
QY 251 TGAACARTGCYACCGAYACNAARAARTTYTNTGWSNYNTTYGCNCGNTGY 310
Db 61 TCGATGACCTAGACGAGACCATCGACCATCGCTCTCGGTGCGAGTGAAGACC 120
QY 311 TNGAYGAYTNGAYGARACNATNCARCNCTGYCAYWSNTTYGTCNCARGTNAARGAYM 370
Db 121 GCTGCGCCCGGTATGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTG 180
QY 371 GNTGCGCCNCGNATGWSNCGNTTYGNTTYCCTGCGCNG-AYATGTYNGARTGYGAY 429
Db 181 CGTTTCCCGGAGGACACACCTTTGATCCCTCGCTAGCAGCGACACCTCTCC 240
QY 430 MGNTTYCCNARGAYAYGAYTNGATNCCNTGNCNNSNNGAYCAYTYNTNCCN 489
Db 241 GCACCCAGGAGCTCAAGAGTATGTGAGCTCGCTGCAAAATAAATGATGATGACAC 300
QY 490 GCNACNARGARGCNCNCAAGTNTGARGCNTGYAARAAYAAARAYGAYGAYAA 549
Db 301 GACATTATGAGACGCTTTGTAAAATGATTTTGCACTTGAATAATNAAGTGAAGGA 360
QY 550 GAYATNA-TGAGACNCTNTGYAARAAYGAYTYGCTN-ARATN-AARGTNAARG-A 605
Db 361 GATAACCT 368

QY 606 RATNACNT 613
: || || |

RESULT 9

LOCUS AA330758 302 bp mRNA EST 21-APR-1997
DEFINITION EST34493 Embryo, 6 week I Homo sapiens cDNA 5' end, mRNA sequence.

ACCESSION AA330758

NID 91983000

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

Eukaryotes; mitochondrial eukaryotes; Metazoa; Chordata;

Vertebrata; Mammalia; Euthera; Primates; Catarrhini; Homiidae;

Homo.

REFERENCE 1 (bases 1 to 302)

AUTHORS

Adams, M.D., Kerlavage, A.R., Fleischmann, R.D., Fuldner, R.A., Bult, C.J., Lee, N.H., Kirkness, E.F., Weinstock, K.G., Gocayne, J.D., White, O., Sutton, G., Blake, J.A., Brandon, R.C., Man, Wai, C., Clayton, R.A., Cline, T.R., Cotton, M.D., Earle-Hughes, J., Fine, L.D., Fitzgerald, L.M., Fitzhugh, W.M., Fritchman, J.L., Geoghegan, N.S., Glodek, A., Gnehm, C.L., Hanna, M.C., Hedblom, E., Hinkle, P.S., Jr., Kelley, J.M., Kelley, J.C., Liu, L.-I., Marmaros, S.M., Merrick, J.M., Moreno-Palauques, R.F., McDonald, L.A., Nguyen, D.T., Peilgrino, S.M., Phillips, C.A., Ryder, S.E., Scott, J.L., Saudek, D.M., Shirley, R., Small, K.V., Spriggs, T.A., Utterback, T.R., Weidman, J.F., Li, Y., Bednarek, D.P., Cao, L., Cepeda, M.A., Coleman, T.A., Collins, E.J., Dimke, D., Feng, D.-F., Fertie, A., Fischer, C., Hastings, G.A., He, W.W., Hu, J.S., Greene, J.M., Gruber, J., Hudson, P., Kim, A.K., Kozak, D.L., Kunsch, C., Hungjun, J., Li, H., Meissner, P.S., Olsen, H., Raymond, L., Wei, Y.F., Wing, J., Xu, C., Yu, G.L., Ruben, S.M., Dillion, P.J., Fannon, M.R., Rosen, C.A., Haseltine, W.A., Fields, C., Fraser, C.M. and Venter, J.C.

Initial assessment of human gene diversity and expression patterns

based upon 83 million nucleotides of cDNA sequence

Nature 377 (6547 Suppl), 3-174 (1995)

Other ESTs: THCL80057

Contact: Kerlavage, AR

COMMENT

Bioinformatics

The Institute for Genomic Research

9712 Medical Center Drive, Rockville, MD 20850 USA

Tel: 3018699056

Fax: 3018699423

Email: arkerlavet@igr.org

For clone availability, additional sequence and expression

information related to this EST, please check the TIGR Human Gene

index (<http://www.tigr.org/tldb/hgi/hgi.html>)

Seq primer: M13 Reverse.

FEATURES

source

1. .302

/organism="Homo sapiens"

/note="Organ: embryo; Vector: pBluescript SK-; Site_1:

EcoRI; Site_2: XhoI"

/db_xref="ATCC (inhost):132379"

/db_xref="taxon:9606"

/clone.lib="Embryo, 6 week I"

/dev_stage="embryo, 6 wks"

<1. .>302

BASE COUNT

99 a 57 c 82 g 56 t 8 others

ORIGIN

Query Match 33.2%; Score 244; DB 4; Length 302;

Best Local Similarity 58.6%; Pred. No. 8.91e-270;

Matches 177; Conservative 69; Mismatches 56; Indels 0; Gaps 0;

Db 1 AACGACATATGAACGCGTTGTAAATAATTTTCTACTGAAATAAAGTGAGGAG 60

QY 547 AAYGAYATNATGARACNTNTGYAARAAYGAYTYGCTYNTAARATNAARGTNAARG 606

Db 61 ATAACCTACATCAACCGAGATACCAATCATCTTGGACCAAGCAAGCAAGCAAGCAAG 120

QY 607 ATNACNTATNAAYMGNGAYACNATNATNTNGARACNARACNARACNATNTAY 666


```
/sex="unknown"  
/dev_stage="12.5dpc total fetus"  
/lab_host="DH10B"  
<1..>394  
mRNA 104 a 96 c 118 g 76 t  
BASE COUNT  
ORIGIN  
  
Query Match 32.0%; Score 235; DB 25; Length 394;  
Best Local Similarity 57.5%; Pred. No. 2,08e-257;  
Matches 164; Conservative 71; Mismatches 50; Indels 0; Gaps 0;  
  
Db 1 AAGGAGATAAGCTACATCAACAGACAGACACCAAGATCATCTCGGAGACAAAGAGACC 60  
QY 601 AARGARATNACNTATYATNAAYMGNGAYACNAARATNATYTNARACNAARACN 660  
  
Db 61 ATTACAGCTGAACGGCGTGTCCGAAAGGACCTGAGAAATCCGTCTGGCTCAAA 120  
QY 661 ATNTAYAAATYATNAAYMGNGAYACNAARATNATYTNARACNAARACN 720  
  
Db 121 GACAGCTGCACTGACCTGTGAGGAGATGAACGACATCAACGCTCCGTATCTGGTCA 180  
QY 721 GAYWNTYNTCACTGACCTGTGAGGAGATGAACGACATCAACGCTCCGTATCTGGTCA 180  
  
Db 181 GGACAGAACGAGCGGGAGCTGTGATGATCACTCCGTGAAACGCTGCGAGAGGCGAG 240  
QY 781 GNCARAAACGAGCGGGAGCTGTGATGATCACTCCGTGAAACGCTGCGAGAGGCGAG 240  
  
Db 241 AGAGAGTTCAAGCGGATCTCCGAGAGCTCCGAGCTGCGAATGC 285  
QY 841 MNGARTTYAARMGNATNWSNMGNSNATNMGNAATYNTCAATGY 885  
  
RESULT 12  
LOCUS W08345 299 bp mRNA EST 05-SEP-1996  
DEFINITION mb41f02.r1 Soares mouse p3NMF19.5 Mus musculus cDNA clone 331995  
ACCESSION W08345  
NID g1282366  
KEYWORDS EST.  
SOURCE house mouse.  
ORGANISM Mus musculus.  
REFERENCE 1 (bases 1 to 299)  
AUTHORS Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T., Geisler,S., Kucaba,I., Lacy,M., Le,M., Martin,J., Morris,M., Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B., Theising,B., Wylie,T., Lennon,G., Soares,B., Willson,R. and Waterston,R.  
TITLE The WashU-HMI Mouse EST Project  
JOURNAL Unpublished (1996)  
COMMENT Contact: Marra M/Mouse EST Project  
WashU-HMI Mouse EST Project  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: mouseest@watson.wustl.edu  
This clone is available royalty-free through LNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.  
MGI:213395  
Seq primer: mob.REGA-ET  
High quality sequence stop: 280.  
Location/Qualifiers  
1..299  
/organism="Mus musculus"  
/note="Vector: pT73D (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was primed with a Not I oligo(dT) primer [5', TGTTCAATCTGAAGTGGAGCGCGCATTTTTTTTTTTT 3'], RI double-stranded cDNA was size selected, ligated to Eco RI
```

```
adapters (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of a modified pT73 vector (Pharmacia). Library went through one round of normalization to a Cot = 5. Library constructed by Bento Soares and M.Fatima Bonaldo. RNA was kindly provided by Dr. Minoru Ko (Wayne State University)."  
/db_xref="taxon:10090"  
/clone_lib="331995"  
/clone_lib="Soares mouse p3NMF19.5"  
/dev_stage="19.5 dpc total fetus"  
/lab_host="DH10B (ampicillin resistant)"  
<1..>299  
mRNA 79 a 95 c 72 g 53 t  
BASE COUNT  
ORIGIN  
  
Query Match 28.6%; Score 210; DB 22; Length 299;  
Best Local Similarity 59.5%; Pred. No. 3.11e-223;  
Matches 175; Conservative 68; Mismatches 46; Indels 5; Gaps 5;  
  
Db 11 GCCTGTCACCTGCTCT-CGTGCAG-TGAAGGACCGCTGCGC-CCGTCATGTCCGCTT 67  
QY 336 RCCNTGYCAWNTYNTGYGTNCARGTNAARGAYMGNTGYCNCNCNTNATGWSNGCNT 395  
  
Db 68 CG-CTTCCCTGGCCAGACATGCTGGAGTGGACCGTTTCCCGAG-ACAACGACCTCTG 125  
QY 396 YGNTTYCCNTGGCCNGAYATGTNGARTGYGAYMGNTTYCCNCARGAYAYGAYTNTG 455  
  
Db 126 CATCCCTCGCTAGTAGCGACGACCTCTCCGCGGCACAGAGAACTCCCAAGGTGTG 185  
QY 456 YATNCCNTTNGCNSWNSNGAYCAIYNTYNTCCNCAGARGCNCNARGTNTG 515  
  
Db 186 TGAAGCCTGCAAAACCAAGATGAGGACGACGACATCATCGTGAACACCTTTGTAAAAA 245  
QY 516 YGAGCNGTGAARAAYARAAYGAYGAYGAYATNATGGARACNTYNTGYAARA 575  
  
Db 246 TGACTTCGCTGAAATCAAGTGAAGGAGATACCTAGTACACAGACAC 299  
QY 576 YGATTYGCNTNAATNAARGTNAARGATNACNTATNATNAAYMGNGAYAC 629  
  
RESULT 13  
LOCUS AA297695 248 bp mRNA EST 18-APR-1997  
DEFINITION EST113252 Fetal heart II Homo sapiens cDNA 5' end, mRNA sequence.  
ACCESSION AA297695  
NID g1950029  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.  
REFERENCE 1 (bases 1 to 248)  
AUTHORS Adams,M.D., Kerlavage,A.R., Fleischmann,R.D., Fuldner,R.A., Bult,C.J., Lee,N.H., Kirknes,E.F., Weinstock,K.G., Gocayne,J.D., White,O., Sutton,G., Blake,J.A., Brandon,R.C., Man-Wai,C., Clayton,R.A., Cline,T.R., Cotton,M.D., Earle-Hughes,J., Fine,L.D., Fitzgerald,L.M., Fitzhugh,W.M., Fritchman,J.L., Geoghagen,N.S., Glodke,A., Gnehm,C.L., Hanna,M.C., Hedblom,E., Hinkle,P.S., Jr., Kelley,J.M., Kelley,J.C., Liu,L.-I., Marmaros,S.M., Merrick,J.M., Moreno-Palauques,R.F., McDonald,L.A., Nguyen,D.T., Pelligrino,S.M., Phillips,C.A., Ryder,S.E., Scott,J.L., Saudek,D.M., Shirley,R., Small,K.V., Spriggs,T.A., Utterback,T.R., Weidman,J.F., Li,Y., Bednarek,D.P., Cao,L., Cepeda,M.A., Coleman,T.A., Collins,E.J., Dimke,D., Feng,D.-F., Ferrie,A., Fischer,C., Hastings,G.A., He,W.W., Hu,J.-S., Greene,J.M., Gruber,J., Hudson,P., Kim,A.K., Kozak,D.L., Kunsch,C., Hungjun,J., Li,H., Weissner,P.S., Olsen,H., Raymond,L., Wei,Y.F., Wing,J., Xu,C., Yu,G.L., Ruben,S.M., Dillion,P.J., Fannon,M.R., Rosen,C.A., Haseltine,W.A., Fields,C., Fraser,C.M. and Venter,J.C.  
TITLE Initial assessment of human gene diversity and expression patterns based upon 83 million nucleotides of cDNA sequence  
JOURNAL Nature 377 (5547 Suppl), 3-174 (1995)  
MEDLINE 96026280
```